

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 06:39:41 ; Search time 4117 Seconds
(without alignments)
10157.117 Million cell updates/sec

Title: US-10-659-869A-35
Perfect score: 863
Sequence: 1 gcacgagctctatcacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382.8	44.4	1022	8	PSMYB26
2	251.2	29.1	745	8	GHY554698
3	239.4	27.7	910	8	AY086615 Arabidops
4	236.2	27.4	645	8	AY519632 Arabidops
5	236.2	27.4	645	8	AF175987 Arabidops
6	216.6	25.1	681	8	AY519589 Arabidops
7	216.6	25.1	1027	8	AK118439 Arabidops
8	215	24.9	1033	6	AB005888 Arabidops
9	215	24.9	1033	6	AB005888 Arabidops
10	208	24.1	1127	8	AY088761 Arabidops
11	206.4	23.9	621	8	AY519582 Arabidops
12	204.8	23.7	652	8	BT005574 Arabidops
13	204.8	23.7	894	8	AK118091 Arabidops
14	194.6	22.5	872	8	AB058642 Lilium hy
15	190	22.0	649	8	AF485900 Dendrobii
16	176.6	20.5	1117	8	AF510112 Craterost
17	175	20.3	1295	8	CPU33917 Arabidops
18	170.6	19.8	861	8	AY519568 Arabidops
19	169.4	19.6	924	8	AY519635 Arabidops

20	169.4	19.6	1347	6	AR439820	Sequence
21	167.8	19.4	822	6	AX507678	Sequence
22	167.8	19.4	822	6	AX507678	Sequence
23	167.8	19.4	822	8	AY519579	Arabidops
24	167.8	19.4	822	8	BT011656	Arabidops
25	167.8	19.4	987	8	BT010949	Arabidops
26	163.4	18.9	972	8	AY519584	Arabidops
27	163.4	18.9	1423	8	AF262733	Arabidops
28	163	18.9	1077	6	AX653037	Sequence
29	161.4	18.7	990	6	AX654067	Sequence
30	161.4	18.7	990	6	AX699688	Sequence
31	161.4	18.7	990	6	AX699719	Sequence
32	161.4	18.7	1575	8	AK107424	Oryza sat
33	161.4	18.7	1645	8	AK120551	Oryza sat
34	161	18.7	1129	8	BT009018	Triticum
35	158.2	18.3	945	6	AX654022	Sequence
36	156.8	18.2	1421	8	AK103455	Oryza sat
37	156.4	18.1	858	6	AX653047	Sequence
38	156.4	18.1	1343	8	AY026332	Oryza sat
39	155.6	18.0	852	8	AY519558	Arabidops
40	155.6	18.0	1140	8	AF334815	Arabidops
41	154.8	17.9	1404	8	AK069082	Oryza sat
42	154	17.8	786	6	AX507599	Sequence
43	154	17.8	786	8	AY519606	Arabidops
44	154	17.8	817	8	AY133705	Arabidops
45	150.4	17.4	903	6	AX653670	Sequence

ALIGNMENTS

RESULT 1

PSMYB26

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

PSMYB26 1022 bp mRNA linear PLN 09-FEB-1998
P. sativum mRNA for Myb-like protein (Myb26).
Y11105
Y11105.1 GI:1841474
Myb-like protein; Myb26.
Pisum sativum (pea)
Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.

Uimari, A. and Strommer, J.
Myb26: a MYB-like protein of pea flowers with affinity for promoters of phenylpropanoid genes
Plant J. 12 (6), 1273-1284 (1997)
98112025
PUBMED 9450341
2 (bases 1 to 1022)
Strommer, J.N.
Direct Submission
Submitted (05-FEB-1997) J.N. Strommer, University Of Guelph,
Molecular Biology And Genetics, Guelph, Ontario N1G 2W1, CANADA
Location/Qualifiers
1. .1022
/organism="Pisum sativum"
/mol_type="mRNA"
/db_xref="taxon:3888"
/tissue_type="flower bud"
92. .745
/codon_start=1
/product="Myb26"
/protein_id="CAA71992.1"
/db_xref="GI:1841475"
/db_xref="GOA:P93474"
/db_xref="UniProt/TrEMBL:P93474"
/translation="MDKKPCNSSQDPEVRKGPWTMEEDLILINYIAHGGVWMSLAK
AAGLKTKGSKRLRLNLPDVRGNITPEOLLIMELSHKGNRWSKAKLHPGRT
DNELKQFWTRIQKHLKQVDNPNQNPQKMSLEINDHHHHHPQSSSVSNLVEPM
ETYSYSYQGTLEFPFFPTQFTTINDDHQNCCANNNNNNYMSMEDISWQLNGD"

ORIGIN

Query Match	44.4%	Score 382.8	DB 8	Length 1022
Best Local Similarity	76.8%	Pred. No. 1.7e-72		
Matches 522	Conservative 0	Mismatches 122	Indels 36	Gaps 3
Qy	8	CTCTATCACACACAAGTCAATGGATAAAAAACAAGTGTAAGACGTCTCAAGATCCT	67	
Db	68	CTCACTCATATACATGCAAAATATATGGACAAAACCCCTGCAACTCATCTCAAGATCCT	127	
Qy	68	GAAGTGAGAAAAGGGCCCTTGGACAATGGAAGAAGACTTTGATCTTGATGAACATATATTGCA	127	
Db	128	GAAGTGAGAAAAGGGCCATGGACCATGGAAGAAGACTTGATTTTGATCAATATATTGCA	187	
Qy	128	AATCATGGGGAAGGTGTTTGGAACTCTTTTGGCCAAAGCTGCTGGTCTCAAAACGTAAACGGA	187	
Db	188	AATCATGGTGAAGGTGTTTGGAAATTCCTTAGCCAAAGCTGCTGTTTAAACGACGCGGA	247	
Qy	188	AAGAGTTTCGGGCTTAAGGTGGCTAAATTTACCTCCGTCCTGATGTTAGAAAGAGGAATATT	247	
Db	248	AAGAGTTTGCAGGCTTCGATGGTTTAAACTATCTTCGTCGGGATGTTAGAAAGAGGAATATT	307	
Qy	248	ACACCCGAGGAACAACCTTTTGATTTATGGAGCTCCACGCAAAAGTGGGAAAAAGGTGGTCC	307	
Db	308	ACACCTGAGGAACAACCTTTTGATCATGGAACTTCATTCAAAGTGGGAAATAGGTGGTCC	367	
Qy	308	AAAATTGCCAAGCATCTACCTGGGAAGGACTGATTAATGAGATCAAGAACTATTGGAGGACA	367	
Db	368	AAAATTGCAAGCATCTTCCAGGAAGAACTGACAAATGAGATTAAGAAATTTTGGAGGACT	427	
Qy	368	AGGATCAGAGGACATCAAGCAAGCTGAGAACTTTCAGCAACAGAGTAGTAATAA----	423	
Db	428	AGGATACAAAGGACATTAAGCAAGTTGATTAACCTTAATCAACAAAATTTTCAGCAAAA	487	
Qy	424	-----TTCTGAGATAAATGATCAC-----CAAGCTAGCACTAGCCAT	460	
Db	488	ATGAGTTTAGAGATAAATGATCATCAACATCATCTCATCAACCAAGTAGTAGCAA	547	
Qy	461	GTPTTCCACCATGGCTGAGCCCATGGAGATGTAATCTCCACCTGTATCAAGGAATGTTA	520	
Db	548	GTPTTCAATTTGGTAGAGCCAATGGAAACTTATTCTCCAACTTCATATCAAGGAACCTTG	607	
Qy	521	GAGCCATTTCAACTCAGTTCCTTACAAATTAATCCTGATC-----AATCCAGTTGT	571	
Db	608	GAGCCATTTCCACTCAATTTCCAACATTAATTAATGATCATCATCAAACTTCAAAATGT	667	
Qy	572	TGTACCAATGACAAACAACATTAATTTGGAGCATGGAGGATAGCTGGTCAATGCAA	631	
Db	668	TGTCCCAATGACAAACAATAACAATTAATTTGGAGCATGGAGGATATCTGGTCAATGCAA	727	
Qy	632	TTACTGAACGGTGATTTAAAT	651	
Db	728	TTACTCAATGGAGATTTAAAT	747	

RESULT 2	
GHY554698	
LOCUS	GHY554698 745 bp mRNA linear PLN 07-JAN-2004
DEFINITION	Gerbera hybrid cv. 'Terra Regina' mRNA for MYB8 protein.
ACCESSION	AJ554698
VERSION	AJ554698.1 GI:40643881
KEYWORDS	myb8 gene; MYB8 protein.
SOURCE	Gerbera hybrid cv. 'Terra Regina'
ORGANISM	Gerbera hybrid cv. 'Terra Regina'
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons; Core eudicots; Asterales; Campanulids; Asterales; Asteraceae; Mutisioideae; Mutisiaeae; Gerbera.
REFERENCE	1 Elomaa,P., Uimari,A., Mehto,M., Albert.V.A., Laitinen,R.A. and Teeri,T.H. Activation of anthocyanin biosynthesis in Gerbera hybrida (Asteraceae) suggests conserved protein-protein and

JOURNAL
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.	1. High Accuracy: The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.
2. Scalability: The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.	2. Scalability: The model is designed to handle large-scale data and complex tasks, making it suitable for enterprise-level applications.
3. Interpretability: The model provides clear and understandable explanations for its predictions, enhancing transparency and trust.	3. Interpretability: The model provides clear and understandable explanations for its predictions, enhancing transparency and trust.
4. Robustness: The model is highly robust to noise and outliers, ensuring reliable performance in real-world scenarios.	4. Robustness: The model is highly robust to noise and outliers, ensuring reliable performance in real-world scenarios.
5. Efficiency: The model is optimized for fast processing and low resource consumption, making it ideal for resource-constrained environments.	5. Efficiency: The model is optimized for fast processing and low resource consumption, making it ideal for resource-constrained environments.

gene
CDS

ORIGIN

Query Match	29.1%;	Score	251.2;	DB	8;	Length	745;
Best Local Similarity	80.2%;	Pred.	No. 5.4e-44;				
Matches	295;	Conservative	0;	Mismatches	73;	Indels	0;
Gaps	0;						
QY	36	AAAAACAACAGTGTAAAGACGTCTCAAGATCCTCAAGCTGAGAAAAAGGCGCTTGGACAAATGG	95				
DB	44	AAAAAAGGCGCAACAATTTTCGCACAGACGTGCAAGTTAGAAAAAGGCCATGACGATGG	103				
QY	96	AGAAGAAGCTTGATCTTTGATGAACTATATTTGCAAAATCATGGGGAAGGTGTTTGGAACTCTTT	155				
DB	104	AGAAGAAGCTTGATTTCTTATAAACTATCATGCGAAATCATGCGAAAGGTGTTTGGAACTCTC	163				
QY	156	TGCGCAAAGCTGCTGCTCTCAACAGTAAAGGAAAGAGTTGCCGGCTTAAGTGTGCTAAATT	215				
DB	164	TTGCCAGATCCGAGGTCTTAAAGAAACCGGAAAAAGTTGCCGGCTTCGGTGGCTAAACT	223				
QY	216	ACCTCCGCTCTGATGTTTAGAAGAGGGAATATTACACCCGAGGAACAACTTTTGATTATGG	275				
DB	224	ATCTCCGGCCAGATGTAAGGAGGGTAAATTATACACCCGAAAGAAACAACTCTTCATCATGG	283				
QY	276	AGCTCAACGCAAAAGTCGGGAAACAGGTGGTCCAAAATTGCCAAGCATCTACCTTGGAGGA	335				
DB	284	AGCTTCATGCCAAATCGGGAAAATAGATGTTCAAAAATAGCCAAACATCTTCCAGGAAGAA	343				
QY	336	CTCATATAGATCAAGAAGCTATTGTGAGGACAAAGGATCCAGAAGCACATCAAGCAAGCTG	395				
DB	344	CAGATACGAGATAAAGAAATTATTGGAGAACAGAATTCAAAAACACATCAAGCAAGTTG	403				
QY	396	AGAAGCTTT	403				
DB	404	AAAAATTTT	411				

RESULT 3
AY086615
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

protein-promoter interactions between the anciently diverged
monocots and eudicots
Plant Physiol. 133 (4), 1831-1842 (2003)
14605235
2 (bases 1 to 745)
Elomaa,P.
Direct Submission
Submitted (28-MAR-2003) Elomaa P., Department of Applied Bio-
University of Helsinki, P.O.Box 27, 00014 University of Helsi
FINLAND

Location/Qualifiers

```

1. .745
/organism="Gerbera hybrid cv. 'Terra Regina'."
/mol_type="mRNA"
/cultivar="Terra Regina"
/db_xref="taxon:22891"
/tissue_type="petal"
1. .745
/gene="myb8"
40. .612
/gene="myb8"
/function="transcription factor"
/codon_start=1
/product="MYB8 protein"
/protein_id="CAD87008.1"
/db_xref="GI:40643882"
/translat="MEKRPNNSSQDVBVRGKPTMEEDLILINVI
SAGKRTGKCLRNLKPLNLRDVRGRNI TPEQLLI MELHAKWGN
DNETKNYWRPTI QKHIVQENFSQTSGGNDQATMSSTQCNV
ENESGPEPTTENNWMSMEDLWSMQLNSR"

```

AY086615 910 bp mRNA linear PLN 14-APR-2003
 Arabidopsis thaliana clone 262460 mRNA, complete sequence.
 AY086615
 AY086615.1 GI:21405325
 FLI CDNA.
 Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 910)
AUTHORS Haas,B.J., Volkovskiy,N., Town,C.D., Troupkan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376

REFERENCE 2 (bases 1 to 910)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 910)
AUTHORS Brover,V., Troupkan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Llaer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES
source
1..910
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="262460"
110..754
/codon_start=1
/product="putative transcription factor MYB24"
/protein_id="AAM63674.1"
/db_xref="GI:21554734"
/translation="MEKRESSGGSGDAEVRKGPWTMEEDLILINVIANHGEVWNS LAKSAGLRKTKSCRLWNLRLPDRVRGNITPEQLTIMELHAKGNRWSKIAKHLPL GRDNEIKNFWRTKIQYIIKSGETTTVGSQSEFINHATTSHTMNDTQTDMDMYSPTTSYQHASNINQNLNYGNYVPESGIMMPLSDVQSEQYNSVDDLWPMNINYN"

CDS
1..910
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
110..754
/codon_start=1
/product="putative transcription factor MYB24"
/protein_id="AAM63674.1"
/db_xref="GI:21554734"
/translation="MEKRESSGGSGDAEVRKGPWTMEEDLILINVIANHGEVWNS LAKSAGLRKTKSCRLWNLRLPDRVRGNITPEQLTIMELHAKGNRWSKIAKHLPL GRDNEIKNFWRTKIQYIIKSGETTTVGSQSEFINHATTSHTMNDTQTDMDMYSPTTSYQHASNINQNLNYGNYVPESGIMMPLSDVQSEQYNSVDDLWPMNINYN"

ORIGIN
Query Match 27.7%; Score 239.4; DB 8; Length 910;
Best Local Similarity 72.4%; Pred. No. 1.9e-41;
Matches 325; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
56 TCTCAAGATCTCTGAAGTGAGAAAGGGCCCTTGGACAATGGAAGAGACTTGTATCTTGTATG 115
143 TCAGGAGATGAGAGTGAGAAAGGGCCATGGACGATGGAAGAGATTTGATTTCTCATC 202
116 AACTATATTGCAATCATGGGGAAGGTGTTTGGAACTCTTTGGCAAAAGCTGCTGGTCTC 175
203 AATTATATCGCAATCATGGTGAAGGTGTTTGGAACTCTCTCGCCAAATCTGCAGGACTA 262
176 AAACCTACGGAAGAGCTTCCGGCTAAGTGGCTTAATTAATTCCTCCGCTGATGTTAGA 235
263 AAACGACCGGGAAGGTGCGGCTCCGGTGGCTGAACTACCTCCGACCTGATGTCGGA 322
236 AGAGGGAATATTACACCCGAGGAACAACACTTTTATTGATGGAGCTCCACGCAAGTGGGGA 295

Db 323 CGGGGAAATATCACACGAGAAAGACAGCTCACCATCATGGAACCTTCATGCAAAATGGGGA 382
Qy 296 AACAGGTGTCCTCAAAATTTGCCAGCATCTACCTGGAGGACTGATATGAGATCAAGAAC 355
Db 383 AATAGGTGGTCAAAATTTGCAAGCATTTACCAGGAAGGCCGCAATGAGATTAAGAAC 442
Qy 356 TATTGGAGGCAAGGATCCAGAGCACATCAAGCAAGCTTGAGAACTTTTCAGCAACAGAGT 415
Db 443 TTTTGGAGGACTAAGATCCAGAAATATCATCATCAAGAGCGGAGAAACGACGACCTTGA 502
Qy 416 AGTAATTAATTTCTGAGATAAATGATCACAAGCTAGCACTAGCCATGTTCCACCATGGCT 475
Db 503 TCACAAAGCTCCGAGTTTATAAACCATCATCGCAACACGAGCCATGTGATGATGATACT 562
Qy 476 --GAGCCCATGGAGATGTTTCTCCACC 501
Db 563 CAAGAAACCATGGATATGTTTCTCCAAC 591

RESULT 4
AY519632
LOCUS Arabidopsis thaliana MYB transcription factor (At5g40350) mRNA, complete cds.
ACCESSION AY519632
VERSION AY519632.1 GI:41619435
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 645)
AUTHORS Qu,L. and Gu,H.
TITLE The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 645)
AUTHORS Qu,L. and Gu,H.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

FEATURES
Location/Qualifiers
1..645
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..645
/locus_tag="At5g40350"
1..645
/locus_tag="At5g40350"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AAS10102.1"
/db_xref="GI:41619436"
/translation="MEKRESSGGSGDAEVRKGPWTMEEDLILINVIANHGEVWNS LAKSAGLRKTKSCRLWNLRLPDRVRGNITPEQLTIMELHAKGNRWSKIAKHLPL GRDNEIKNFWRTKIQYIIKSGETTTVGSQSEFINHATTSHTMNDTQTDMDMYSPTTSYQHASNINQNLNYGNYVPESGIMMPLSDVQSEQYNSVDDLWPMNINYN"

ORIGIN
Query Match 27.4%; Score 236.2; DB 8; Length 645;
Best Local Similarity 71.9%; Pred. No. 9.7e-41;
Matches 323; Conservative 0; Mismatches 123; Indels 3; Gaps 1;
56 TCTCAAGATCTCTGAAGTGAGAAAGGGCCCTTGGACAATGGAAGAGACTTGTATCTTGTATG 115
Db 34 TCAGGAGATGAGAGTGAGAAAGGGCCATGGACGATGGAAGAGATTTGATTTCTCATC 93
Qy 116 AACTATATTGCAATCATGGGGAAGGTGTTTGGAACTCTTTGGCAAAAGCTGCTGGTCTC 175

```
Db 94 AATTATATCGCCCAATCATGGTGAAGGTGTTTGGAACTCTCTCGCCAAATCTCGAGGACTA 153
QY 176 AAACCTAACGGAAGAGTTCGCGGCTAAGGTGGCTAAATTAACCTCCGCTCGATGTTAGA 235
Db 154 AAACGACACGGGAAAAGTTGCGGCTCGGTGGCTGAACTACCTCCGACCTGATGTGCGA 213
QY 236 AGAGGGAATATTACACCCGAGGAAACAATTTTGTATTATGGAGCTCCACGCAAAAGTGGGGA 295
Db 214 CGGGGAATATACACACGAGAGACAGCTCACCATCATGGAACTTTCATGCAAAATGGGGA 273
QY 296 AACAGGTGGTCAAAATTTGCCAAGCATCTACCTGGAAGAGCTGATTAATGAGATCAAGAAC 355
Db 274 AATAGTGGTCAAAATTTGCAAGCATTTTACCAGGAGAACCGACAATGAGATAAGAAAT 333
QY 356 TATTCGAGSACAGCATCCAGAGCACATCAGCAAGCTCAGCAAGCTTTTCAGCAACAGAGT 415
Db 334 TTTTGGAGACTTAAGATCCAGAAATATACATCATCAAGCGGAGAAACAGCACCGTTGGA 393
QY 416 AGTAATAATTCTGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCT 475
Db 394 TCACAAAGCTCGAGTTTATATAACCATCATCGACACGAGCCATGTCATGAATGATACT 453
QY 476 ---GAGCCCATGAGATGATTTCTCCACC 501
Db 454 CAAGAAACCATGGATATGTTTCTCCAAC 482
```

```
RESULT 5
AF175987 645 bp mRNA linear PLN 30-AUG-2001
LOCUS Arabidopsis thaliana putative transcription factor (MYB24) mRNA,
DEFINITION complete cds.
ACCESSION AF175987 GI:5823306
VERSION AF175987.1
KEYWORDS ORGANISM
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 645)
Kranz,H.D., Denekamp,M., Greco,R., Jin,H., Leyva,A., Meissner,R.C.,
Petróni,K., Urzainqui,A., Bevan,M., Martin,C., Smeekens,S.,
Tonelli,C., Paz-Ares,J. and Weisshaar,B.
Towards functional characterisation of the members of the R2R3-MYB
gene family from Arabidopsis thaliana
Plant J. 16 (2), 263-276 (1998)
```

```
JOURNAL MEDLINE 99056848
PUBMED 9839469
REFERENCE 2 (bases 1 to 645)
AUTHORS Stracke,R., Werber,M. and Weisshaar,B.
TITLE The R2R3-MYB gene family in Arabidopsis thaliana
JOURNAL Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)
MEDLINE 21481677
PUBMED 11597504
REFERENCE 3 (bases 1 to 645)
AUTHORS Stracke,R. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1999) Dept. Biochemie, Max-Planck-Institut fuer
Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
FEATURES
source
1..645
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Columbia"
1..645
/gene="MYB24"
1..645
/gene="MYB24"
/note="R2R3-MYB family member"
/codon_start=1
/product="putative transcription factor"
```

```
/protein_id="AAD53092.1"
/db_xref="GI:5823307"
/translation="MEKRSSGSGSGDAEVRKGPWTMEEDLILINYIANHGEVWNS
LAKSAGLKRTGKSCRLWLYRPDVRGNITPEEOLTIMELHAKGNRWSKIAKHL
RTDNEIKNFWRTKIQTIIKSGSTTVGSSQSEFINHATTSFMDTQTMDMYS
TTSYQHSNINQNLNYGNYPESGSIIMPLVSVDSEQNWSVDLWPMNIYNGN"
ORIGIN
Query Match 27.4%; Score 236.2; DB 8; Length 645;
Best Local Similarity 71.9%; Pred. No. 9.7e-41;
Matches 323; Conservative 0; Mismatches 123; Indels 3; Gaps 1;
QY 56 TCTCAAGATCTCTGAAGTGAGAAAAAGGCGCTTTGGCAATGGAAGAGACTTGTGATCTTGATG 115
Db 34 TCAGGAGATGCAGAGGTGAGAAAAAGGCGCATGGACGATGGAAGAAGATTGATTCTCATC 93
QY 116 AACTATATTGCAATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTC 175
Db 94 AATTATATCGCCAATCATGGTGAAGGTGTTTGGAACTCTCTCGCCAAATCTCGAGGACTA 153
QY 176 AAACGTAACGGAAGAGTTGCCGCTAAGGTGGCTAAATTAACCTCCGCTCGATGTTAGA 235
Db 154 AAACGACCGGGAAGAGTTGCCGCTCCGTTGGCTGAACCTCCGACCTGATGTGCGA 213
QY 236 AGAGGGAATATTACACCCGAGGAAACAATTTTGTATTATGGAGCTCCACGCAAAAGTGGGA 295
Db 214 CGGGGAATATACACACGAGAGACAGCTCACCATCATGGAACCTTCATGCAAAATGGGGA 273
QY 296 AACAGGTGGTCAAAATTTGCCAAGCATCTACTGGAAGAGCTGATTAATGAGATCAAGAAC 355
Db 274 AATAGTGGTCAAAATTTGCAAGCATTTTATAAACCATCATCGACACGAGCCATGTCATGAATGATACT 453
QY 476 ---GAGCCCATGAGATGATTTCTCCACC 501
Db 454 CAAGAAACCATGGATATGTTTCTCCAAC 482
RESULT 6
AF175989 681 bp mRNA linear PLN 07-FEB-2004
LOCUS Arabidopsis thaliana MYB transcription factor (At3g27810) mRNA,
DEFINITION complete cds.
ACCESSION AF175989
VERSION AF175989.1 GI:41619251
KEYWORDS ORGANISM
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 681)
QY 116 AACTATATTGCAATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTC 175
Db 94 AATTATATCGCCAATCATGGTGAAGGTGTTTGGAACTCTCTCGCCAAATCTCGAGGACTA 153
QY 176 AAACGTAACGGAAGAGTTGCCGCTAAGGTGGCTAAATTAACCTCCGCTCGATGTTAGA 235
Db 154 AAACGACCGGGAAGAGTTGCCGCTCCGTTGGCTGAACCTCCGACCTGATGTGCGA 213
QY 236 AGAGGGAATATTACACCCGAGGAAACAATTTTGTATTATGGAGCTCCACGCAAAAGTGGGA 295
Db 214 CGGGGAATATACACACGAGAGACAGCTCACCATCATGGAACCTTCATGCAAAATGGGGA 273
QY 296 AACAGGTGGTCAAAATTTGCCAAGCATCTACTGGAAGAGCTGATTAATGAGATCAAGAAC 355
Db 274 AATAGTGGTCAAAATTTGCAAGCATTTTATAAACCATCATCGACACGAGCCATGTCATGAATGATACT 453
QY 476 ---GAGCCCATGAGATGATTTCTCCACC 501
Db 454 CAAGAAACCATGGATATGTTTCTCCAAC 482
REFERENCE
AUTHORS The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide
Cloning and Expression Pattern Analysis
TITLE Unpublished
JOURNAL 2 (bases 1 to 681)
REFERENCE Qu,L. and Gu,H.
AUTHORS Direct Submission
TITLE Submitted (07-JAN-2004) Life Sciences, National Laboratory of
JOURNAL Protein Engineering and Plant Genetic Engineering, Peking
University, Beijing 100871, China
FEATURES
source
1..681
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```



```

1. .681
/db_xref="taxon:3702"
/locus_tag="At3g27810"
1. .681
/locus_tag="At3g27810"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AAS10059.1"
/db_xref="GI:41619252"
/translation="MEKRGSGSGSSGSSAEAEVRKGPWTMEEDLILINYIANHGDGV
WNSLAKGAGLRTGKSCRLRWLNPDPVRRGNTTPEOLIMELHAKGNRWKSIK
HLPGRTNEIKFWRTRIOKYIKOSDVTTSSVSGSHSSSEINDOAASTSHNVFCTOD
QAMETYSPTFSTYQHTNMEFNYSAAAVTATVDYVPVMTVDQGTGENTWGMDDIWS
SMHLINGN"

ORIGIN
Query Match 25.1%; Score 216.6; DB 8; Length 681;
Best Local Similarity 77.0%; Pred. No. 1.7e-36;
Matches 264; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 54 CGTCTCAAGATCCTGAAGTGAGAAAGGCGCTTGGCAATGGAAGAGACTTGATCTTGA 113
Db 41 CATCAGCAGAGCAGAGAGTGAGAAAGGACCATGAGCAGATGGAAGAGATCTTATCTTA 100

Qy 114 TGAATATATGCGAAATCATCGGAGAGTGTTTGGAACTCTTTGGCCAAAGCTGCTGTC 173
Db 101 TCAACTATATCGCCAAACACGCGGATGTTGTTGGAAATCTCTCGCCAAATCTGCAGTC 160

Qy 174 TCAACAGTAAACGGAAGTTGCCGCTAAAGTGGCTAAATTAACCTCGTCTCATGTTA 233
Db 161 TAAACGAGACCGGAAAGTTGCCGCTCCGGTGGCTGAATCTCTCGCCCGACGTC 220

Qy 234 GAAGAGGGAATATTACACCGGAGGAACAATTTTGTATTGAGCTCCACGCAAGTGGG 293
Db 221 GACGGGAAACATCACTCCAGAGAGCAATTTATCATCATGGAATCTCATGTAAGTGG 280

Qy 294 GAAACAGTGTCCAAATTCGCAAGCATCTACCTGAGAGACTGATAATGAGATCAAGA 353
Db 281 GAAACAGTGTGCGAAATTCGCAACATCTTCCAGGAAGACGCAACAGATCAAAA 340

Qy 354 ACTATTGAGGACAGAGATCCAGAGCAGCATCAAGCAAGCTGA 396
Db 341 ATTTCTGGAGGACAGAAATTCAAAATATACATCAAGCAATCGGA 383

RESULT 7
AK118439
LOCUS Arabidopsis thaliana At3g27810 mRNA for putative myb family
transcription factor (Atmyb3), complete cds, clone: RAFL19-68-M03.
AK118439
AK118439.1 GI:26451910
FLI CDNA: CAP trapper.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Published Only in Database (2002)
2
(bases 1 to 1027)
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail: meeki@gs.riken.go.jp,
URL: http://pfigweb.gsc.riken.go.jp, Tel: 81-45-503-9625,

COMMENT
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda ELC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified pBluescript vector.
Please visit our web site (http://pfigweb.gsc.riken.go.jp/) for
further details.
Location/Qualifiers
source 1. 1027
/organism="Arabidopsis thaliana"
/mol_type="mrna"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL19-68-M03"
/ecotype="Columbia"
/notes="Common name: thale cress"
1. 1027
/genes="At3g27810/MGF10_23"
135. .815
/genes="At3g27810/MGF10_23"
/codon_start=1
/product="putative myb family transcription factor Atmyb3"
/protein_id="BAC43048.1"
/db_xref="GI:26451911"
/translation="MEKRGSGSGSSGSSAEAEVRKGPWTMEEDLILINYIANHGDGV
WNSLAKGAGLRTGKSCRLRWLNPDPVRRGNTTPEOLIMELHAKGNRWKSIK
HLPGRTNEIKFWRTRIOKYIKOSDVTTSSVSGSHSSSEINDOAASTSHNVFCTOD
QAMETYSPTFSTYQHTNMEFNYSAAAVTATVDYVPVMTVDQGTGENTWGMDDIWS
SMHLINGN"

FEATURES
source 1. 1027
/organism="Arabidopsis thaliana"
/mol_type="mrna"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL19-68-M03"
/ecotype="Columbia"
/notes="Common name: thale cress"
1. 1027
/genes="At3g27810/MGF10_23"
135. .815
/genes="At3g27810/MGF10_23"
/codon_start=1
/product="putative myb family transcription factor Atmyb3"
/protein_id="BAC43048.1"
/db_xref="GI:26451911"
/translation="MEKRGSGSGSSGSSAEAEVRKGPWTMEEDLILINYIANHGDGV
WNSLAKGAGLRTGKSCRLRWLNPDPVRRGNTTPEOLIMELHAKGNRWKSIK
HLPGRTNEIKFWRTRIOKYIKOSDVTTSSVSGSHSSSEINDOAASTSHNVFCTOD
QAMETYSPTFSTYQHTNMEFNYSAAAVTATVDYVPVMTVDQGTGENTWGMDDIWS
SMHLINGN"

ORIGIN
Query Match 25.1%; Score 216.6; DB 8; Length 1027;
Best Local Similarity 77.0%; Pred. No. 1.6e-36;
Matches 264; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 54 CGTCTCAAGATCCTGAAGTGAGAAAGGCGCTTGGCAATGGAAGAGACTTGATCTTGA 113
Db 175 CATCAGCAGAGCAGAGAGTGAGAAAGGACCATGAGCAGATGGAAGAGATCTTATCTTA 234

Qy 114 TGAATATATGCGAAATCATCGGAGAGTGTTTGGAACTCTTTGGCCAAAGCTGCTGTC 173
Db 235 TCAACTATATCGCCAAACACGCGGATGTTGTTGGAAATCTCTCGCCAAATCTGCAGTC 294

Qy 174 TCAACGTAACGGAAGAGTTGCCGCTAAAGTGGCTAAATTAACCTCGTCTCATGTTA 233
Db 295 TAAACGAAACCGGAAAGTTGCCGCTCCGGTGGCTGAATCTCTCGCCCGACGTC 354

Qy 234 GAAGAGGGAATATTACACCGGAGGAACAATTTTGTATTGAGCTCCACGCAAGTGGG 293
Db 355 GACGGGAAACATCACTCCAGAGAGCAACTTATCATCATGGAATCTCATGCTAAGTGGG 414

Qy 294 GAAACAGTGTCCAAATTCGCAAGCATCTACCTGAGAGACTGATAATGAGATCAAGA 353
Db 415 GAAACAGTGTGCGAAATTCGCAACATCTTCCAGGAAGACGCAACAGATCAAAA 474

Qy 354 ACTATTGAGGACAGAGATCCAGAGCAGCATCAAGCAAGCTGA 396
Db 475 ATTTCTGGAGGACAGAAATTCAAAATATACATCAAGCAATCGGA 517

RESULT 8
AR439844
LOCUS Arabidopsis thaliana At3g27810 mRNA for putative myb family
transcription factor (Atmyb3), complete cds, clone: RAFL19-68-M03.
AR439844
AR439844.1 GI:42665809
FLI CDNA: CAP trapper.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Published Only in Database (2002)
2
(bases 1 to 1027)
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail: meeki@gs.riken.go.jp,
URL: http://pfigweb.gsc.riken.go.jp, Tel: 81-45-503-9625,
```

AUTHORS Heard, J., Brown, P., Riechmann, J.L., Keddie, J., Pineda, O., Adam, L., Samaha, R., Zhang, J., Yu, G.-L., Ratcliffe, O., Pilgrim, M., Jiang, C.-Z. and Reuber, L.
TITLE Transgenic plants comprising polynucleotides encoding transcription factors that confer disease tolerance
JOURNAL Patent: US 6664446-A 87 16-DEC-2003;
FEATURES Location/Qualifiers
source 1. .1033
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 24.9%; Score 215; DB 6; Length 1033;
 Best Local Similarity 76.7%; Pred. No. 3.4e-36;
 Matches 263; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 54 CGTCTCAAGATCCTGAAGTGAGAAAGGGCTTGGACAAATGGAAGAACATTTGATCTTGA 113
 Db 155 CATCAGCAGACAGACAGAGTGAGAAAGGACCATGGACGATGGAAGAGATCTTATCTTA 214
 Qy 114 TGAACATATATGCAAAATCATGGGAAGGTGTTTGGAACTTTTGGCCAAAGCTGCTGGTC 173
 Db 215 TCAACTATATCGCCAAACACGCGCATGTTGTTTGGAAATCTCTCCCAAATCTGCAGGTC 274
 Qy 174 TCAAAACGTAAACGGAAGAGTTGCCGGCTAAAGTGGCTAAATTTACCTCCGTCCTGATGTTA 233
 Db 275 TAAACAGAACCGGGAAGAGTTGCCGGCTCCGGTGGCTGAATCTCTCCGCCCGACGTAC 334
 Qy 234 GAAGAGGGAATATTACCCGAGGAACAACTTTTGAATTTGAGCTCCACGCAAGTGGG 293
 Db 335 GACGGGGAACATCACTCCAGAAGAGCAACTTATCATCATGGAATTCATGCTAAGTGGG 394
 Qy 294 GAAACAGGTGGTCCAAATTTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGA 353
 Db 395 GAAACAGGTGGTCCAAATTTGCCAAGCATCTTCCAGGAAGAACGACGACGATCAAAA 454
 Qy 354 ACTATTGGAGGACAGGATCCAGAAGCACATCAAGCAAGCTGA 396
 Db 455 ATTTCTGTAGGACAGAAATTCAAANAATACATCAAGCAATCGA 497

RESULT 9
AB005888 1033 bp mRNA linear PLN 14-FEB-2004
LOCUS Arabidopsis thaliana mRNA for ATMVB3, complete cds.
DEFINITION Arabidopsis thaliana mRNA for ATMVB3, complete cds.
ACCESSION AB005888
VERSION AB005888.1 GI:2280527
KEYWORDS ATMVB3.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Noji, M., Urao, T., Shinozaki, K. Y. and Shinozaki, K.
TITLE Molecular cloning of two cDNAs encoding novel myb homologs from Arabidopsis (Accession Nos. AB005888 and AB005889) (PCR98-111)
JOURNAL Plant Physiol. 117, 720 (1998)
REFERENCE 2 (bases 1 to 1033)
AUTHORS Noji, M.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-1997) Masaaki Noji, Chiba University, Faculty of Pharmaceutical Sciences; Yayoi-cho 1-33, Inage-ku, Chiba, Chiba 263, Japan (E-mail: mnoji@p.chiba-u.ac.jp, Tel: +81-43-290-2906, Fax: +81-43-290-2905)
FEATURES Location/Qualifiers
source 1. .1033
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /ecotype="Columbia"
 1. .1033
 /gene="Atmyb3"

gene

CDS

115. .795
 /gene="Atmyb3"
 /codon_start=1
 /product="ATMYB3"
 /protein_id="BAA21618.1"
 /db_xref="GI:2280528"
 /translation="MEKRGSGSGSGSSAAEVRKGPWTWEEILLIINYIANHGQGV
 WNSLAKSAGLKRKTKSCLRWLNLVDPVRRGNITPEQLIMELHAKWGNRWKSIK
 HLPGRDNEIKFRCRIQKVIKQSDVITTTSSVSGSHSSINDQAASTSSHNVCTQD
 QAMETYSPTPTSYOHTNMEFNYSAAAVTATVDYVPVPTVDDQTCENYMGMDIWS
 SMHLNGN"

ORIGIN

Query Match 24.9%; Score 215; DB 8; Length 1033;
 Best Local Similarity 76.7%; Pred. No. 3.4e-36;
 Matches 263; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 54 CGTCTCAAGATCCTGAAGTGAGAAAGGGCTTGGACAAATGGAAGAACATTTGATCTTGA 113
 Db 155 CATCAGCAGACGACAGAGTGAGAAAGGACCATGGACGATGGAAGAGATCTTATCTTA 214
 Qy 114 TGAACATATATGCAAAATCATGGGAAGGTGTTTGGAACTTTTGGCCAAAGCTGCTGGTC 173
 Db 215 TCAACTATATCGCCAAACACGCGCATGTTGTTTGGAAATCTCTCCCAAATCTGCAGGTC 274
 Qy 174 TCAAAACGTAAACGGAAGAGTTGCCGGCTAAAGTGGCTAAATTTACCTCCGTCCTGATGTTA 233
 Db 275 TAAACAGAACCGGGAAGAGTTGCCGGCTCCGGTGGCTGAATCTCTCCGCCCGACGTAC 334
 Qy 234 GAAGAGGGAATATTACCCGAGGAACAACTTTTGAATTTGAGCTCCACGCAAGTGGG 293
 Db 335 GACGGGGAACATCACTCCAGAAGAGCAACTTATCATCATGGAATTCATGCTAAGTGGG 394
 Qy 294 GAAACAGGTGGTCCAAATTTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGA 353
 Db 395 GAAACAGGTGGTCCAAATTTGCCAAGCATCTTCCAGGAAGAACGACGACGATCAAAA 454
 Qy 354 ACTATTGGAGGACAGGATCCAGAAGCACATCAAGCAAGCTGA 396
 Db 455 ATTTCTGTAGGACAGAAATTCAAANAATACATCAAGCAATCGA 497

RESULT 10
AY088761 1127 bp mRNA linear PLN 14-APR-2003
LOCUS Arabidopsis thaliana clone 94595 mRNA, complete sequence.
DEFINITION Arabidopsis thaliana clone 94595 mRNA, complete sequence.
ACCESSION AY088761
VERSION AY088761.1 GI:21407535
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1127)
AUTHORS Haas, B.J., Volfovsky, N., Town, C.D., Troughan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376
REFERENCE 2 (bases 1 to 1127)
AUTHORS Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1127)
AUTHORS Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Llaer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source
1. .1127
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="94595"
274. .894
/codon_start=1
/product="putative transcription factor"
/protein_id="AAM67076.1"
/db_xref="GI:21618026"
/translation="METTTKKGRVKATITTSQKEEGTVRKGPWTMEEDFILFNYILN HGEGLMSVAKSGLRKTKGKRLRWLYLRPDVRRGNITTEBQLLIIOHAKLGNRW SKIAKHLPGRTDNEIKNFWRTKIQRMKVSNNMNMHHCNSGNSQSSGWTQSSGK AIDTAESFSQAKTTTFNVVEQQSNENYNNVVDLWPHVLLNGDHHVI"

CDS

274. .894
/codon_start=1
/product="putative transcription factor"
/protein_id="AAM67076.1"
/db_xref="GI:21618026"
/translation="METTTKKGRVKATITTSQKEEGTVRKGPWTMEEDFILFNYILN HGEGLMSVAKSGLRKTKGKRLRWLYLRPDVRRGNITTEBQLLIIOHAKLGNRW SKIAKHLPGRTDNEIKNFWRTKIQRMKVSNNMNMHHCNSGNSQSSGWTQSSGK AIDTAESFSQAKTTTFNVVEQQSNENYNNVVDLWPHVLLNGDHHVI"

ORIGIN

Query Match 24.1%; Score 208; DB 8; Length 1127;
Best Local Similarity 77.1%; Pred. No. 1.1e-34;
Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 60 AAGATCTCTGAAGTGAAGAAAGGCGCTTGCAAAATGGAAGAGACTTGATCTTGATGAAC 119
Db 335 AAGAAGGAACAGTGAAGAAAGACCTTGCAATGGAAGAGATTTATCTCTTTTAATT 394
Qy 120 ATATTGCAAAATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAA 179
Db 395 ACATCTTTAATCATGCTGAAGGTCTTTGGAACCTCTGCGCAAAAGCCTCTGCTCTAAAC 454
Qy 180 GTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTGATGTAGAAGAG 239
Db 455 GTACTGGAAGAAAGTGTCTGGCTCGGCTGGCTGAACATCTCTCCGACAGATGTGGCGGAG 514
Qy 240 GGAATATTACACCCGAGGAACAACCTTTTGTATTGAGCTCCAGCAAGTGGGGAACA 299
Db 515 GGAACATTAACACAGAGACAGCTTTTGTATCATTTCACTTCACTTAAGCTTGAACA 574
Qy 300 GGTGTCCTCAAAATTCGCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATT 359
Db 575 GGTGTCGAAGATTCGGAAGCATCTTCGGGAAGAACGACAAACGAGATAAAGAACTTCT 634
Qy 360 GGAGACAAAGATTCAGAGACACATCA 387
Db 635 GGAGACAAAGATTCAGAGACACATGAA 662

RESULT 11

AY519582
LOCUS Arabidopsis thaliana MYB transcription factor (At3g01530) mRNA, complete cds.
DEFINITION
ACCESSION AY519582
VERSION AY519582
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE

AUTHORS Qu,L. and Gu,H.
TITLE The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 621)
AUTHORS Qu,L. and Gu,H.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

FEATURES

source
1. .621
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1. .621
/locus_tag="At3g01530"
1. .621
/locus_tag="At3g01530"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AA510052.1"
/db_xref="GI:41619222"
/translation="METTTKKGRVKATITTSQKEEGTVRKGPWTMEEDFILFNYILN HGEGLMSVAKSGLRKTKGKRLRWLYLRPDVRRGNITTEBQLLIIOHAKLGNRW SKIAKHLPGRTDNEIKNFWRTKIQRMKVSNNMNMHHCNSGNSQSSGWTQSSGK AIDTAESFSQAKTTTFNVVEQQSNENYNNVVDLWPHVLLNGDHHVI"

gene

1. .621
/locus_tag="At3g01530"

CDS

1. .621
/locus_tag="At3g01530"

/codon_start=1
/product="MYB transcription factor"

/protein_id="AA510052.1"

/db_xref="GI:41619222"

60 AAGATCTCTGAAGTGAAGAAAGGCGCTTGCAAAATGGAAGAGACTTGATCTTGATGAAC 119
62 AAGAAGGAACAGTGAAGAAAGACCTTGCAATGGAAGAGATTTATCTCTTTTAATT 121
120 ATATTGCAAAATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAA 179
122 ACATCTTTAATCATGCTGAAGGTCTTTGGAACCTCTGCGCAAAAGCCTCTGCTCTAAAC 181
180 GTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTGATGTAGAAGAG 239
182 GTACTGGAAGAAAGTGTCTGGCTCGGCTGGCTGAACATCTCTCCGACAGATGTGGCGGAG 241
240 GGAATATTACACCCGAGGAACAACCTTTTGTATTGAGCTCCAGCAAGTGGGGAACA 299
242 GGAACATTAACACAGAGACAGCTTTTGTATCATTTCACTTCACTTAAGCTTGAACA 301
300 GGTGTCCTCAAAATTCGCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATT 359
302 GGTGTCGAAGATTCGGAAGCATCTTCGGGAAGAACGACAAACGAGATAAAGAACTTCT 361
360 GGAGACAAAGATTCAGAGACACATCA 387
362 GGAGACAAAGATTCAGAGACACATGAA 389

ORIGIN

Query Match 23.9%; Score 206.4; DB 8; Length 621;
Best Local Similarity 76.8%; Pred. No. 2.7e-34;
Matches 252; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 60 AAGATCTCTGAAGTGAAGAAAGGCGCTTGCAAAATGGAAGAGACTTGATCTTGATGAAC 119
Db 62 AAGAAGGAACAGTGAAGAAAGACCTTGCAATGGAAGAGATTTATCTCTTTTAATT 121
Qy 120 ATATTGCAAAATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAA 179
Db 122 ACATCTTTAATCATGCTGAAGGTCTTTGGAACCTCTGCGCAAAAGCCTCTGCTCTAAAC 181
Qy 180 GTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTGATGTAGAAGAG 239
Db 182 GTACTGGAAGAAAGTGTCTGGCTCGGCTGGCTGAACATCTCTCCGACAGATGTGGCGGAG 241
Qy 240 GGAATATTACACCCGAGGAACAACCTTTTGTATTGAGCTCCAGCAAGTGGGGAACA 299
Db 242 GGAACATTAACACAGAGACAGCTTTTGTATCATTTCACTTCACTTAAGCTTGAACA 301
Qy 300 GGTGTCCTCAAAATTCGCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATT 359
Db 302 GGTGTCGAAGATTCGGAAGCATCTTCGGGAAGAACGACAAACGAGATAAAGAACTTCT 361
Qy 360 GGAGACAAAGATTCAGAGACACATCA 387
Db 362 GGAGACAAAGATTCAGAGACACATGAA 389

RESULT 12

BT005574
LOCUS Arabidopsis thaliana clone U50886 putative myb family transcription factor (At3g01530) mRNA, complete cds.
DEFINITION
ACCESSION BT005574
VERSION BT005574.1
KEYWORDS GI:28973338
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 621)

Qu,L. and Gu,H.

The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis

Unpublished

2 (bases 1 to 621)

Qu,L. and Gu,H.

Direct Submission

Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

Location/Qualifiers

1. .621

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

1. .621

/locus_tag="At3g01530"

1. .621

/locus_tag="At3g01530"

/codon_start=1

/product="MYB transcription factor"

/protein_id="AA510052.1"

/db_xref="GI:41619222"

60 AAGATCTCTGAAGTGAAGAAAGGCGCTTGCAAAATGGAAGAGACTTGATCTTGATGAAC 119
62 AAGAAGGAACAGTGAAGAAAGACCTTGCAATGGAAGAGATTTATCTCTTTTAATT 121
120 ATATTGCAAAATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAA 179
122 ACATCTTTAATCATGCTGAAGGTCTTTGGAACCTCTGCGCAAAAGCCTCTGCTCTAAAC 181
180 GTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTGATGTAGAAGAG 239
182 GTACTGGAAGAAAGTGTCTGGCTCGGCTGGCTGAACATCTCTCCGACAGATGTGGCGGAG 241
240 GGAATATTACACCCGAGGAACAACCTTTTGTATTGAGCTCCAGCAAGTGGGGAACA 299
242 GGAACATTAACACAGAGACAGCTTTTGTATCATTTCACTTCACTTAAGCTTGAACA 301
300 GGTGTCCTCAAAATTCGCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATT 359
302 GGTGTCGAAGATTCGGAAGCATCTTCGGGAAGAACGACAAACGAGATAAAGAACTTCT 361
360 GGAGACAAAGATTCAGAGACACATCA 387
362 GGAGACAAAGATTCAGAGACACATGAA 389

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 652)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones
Unpublished

2 (bases 1 to 652)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers
1. .652
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U50886"
/ecotype="Columbia"
/note="This clone is in pUNI 51."

gene
1. .652
/gene="At3g01530"

CDS
1. .621
/gene="At3g01530"
/codon_start=1
/evidence="experimental"
/product="putative myb family transcription factor"

misc_difference 87
/gene="At3g01530"
/note="compared to genomic sequence"
/replace="t"
622. .652
/gene="At3g01530"

3'UTR
622. .652
/gene="At3g01530"

ORIGIN
Query Match 23.7%; Score 204.8; DB 8; Length 652;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 652)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones
Unpublished

2 (bases 1 to 652)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers
1. .652
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U50886"
/ecotype="Columbia"
/note="This clone is in pUNI 51."

gene
1. .652
/gene="At3g01530"

CDS
1. .621
/gene="At3g01530"
/codon_start=1
/evidence="experimental"
/product="putative myb family transcription factor"

misc_difference 87
/gene="At3g01530"
/note="compared to genomic sequence"
/replace="t"
622. .652
/gene="At3g01530"

3'UTR
622. .652
/gene="At3g01530"

ORIGIN
Query Match 23.7%; Score 204.8; DB 8; Length 652;

Best Local Similarity 76.5%; Pred. No. 6e-34;
Matches 251; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 60 AAGATCTCGAAGTCAGAAAAGGGCTTGGCAATGGAAGAGACTTGATCTTGATGAAC 119
|||||
Db 62 AAGAAGAACAGTGAAGAAAGGACCATGGATATGGAGAAGATTTCATCTCTTTAATT 121
|||||
Qy 120 ATATTGCAATCATCGGGGAAGTGCTTTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAA 179
|||||
Db 122 ACATCTTAATCATGGTGAAGGTCTTTGGAACTCTGTGCGCAAGCCCTCTGGTCTAAAAC 181
|||||

Qy 180 GTAACGGAAGAGTTGCGGCTAAGTGCGCTAAATACCTCCGTCCTGATGTTAGAAGAG 239
|||||
Db 182 GTACTGGAAAAGTTGTGGCTCCGGCTCGGCTGGAATCATCTCCGACCATGTGGCGCGAG 241
|||||

Qy 240 GGAATATTACACCCGAGGAACAACCTTTTGTATTATGGAGCTCCACGAAAGTGGGGAACA 299
|||||
Db 242 GGAACATAACCGAAGAAAGACAGCTTTTGTATCATCTCAGCTTCATGCTTAAGCTTGGAAACA 301
|||||

Qy 300 GGTGGTCCAAATTGCCCAAGCATCTACCTGGAAGGACTGATAATCAGATCAAGAACTATT 359
|||||
Db 302 GGTGGTGAAGTTGCGAAGCATCTTCGGGAAGAACGACACGAGATAAAGAACTTCT 361
|||||

Qy 360 GGAGACAAAGATCCAGAACGACATCAA 387
|||||
Db 362 GGAGACAAAGATTCAGAGACACATGAA 389
|||||

RESULT 13
AKI18091 894 bp mRNA linear PLN 14-FEB-2004
LOCUS Arabidopsis thaliana At3g01530 mRNA for putative transcription factor, complete cds, clone: RAFL19-34-G03.
DEFINITION AKI18091
ACCESSION AKI18091.1 GI:26451233
VERSION FLI_CDNA; CAP trapper.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1
AUTHORS Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
TITLE Arabidopsis thaliana full-length cDNA
JOURNAL Published Only in database (2002)
REFERENCE 2 (bases 1 to 894)
AUTHORS Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:mskic@gsr.riken.go.jp, URL:http://pfweb.gsc.riken.go.jp, Tel:81-45-503-9625, Fax:81-45-503-9586)
COMMENT An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified plasmid vector.
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.
Location/Qualifiers
1. .894
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL19-34-G03"
/ecotype="Columbia"

```

gene
CDS
/note="common name: thale cress"
1. .894
/gene="At3g01530/F4P13_8"
109. .729
/gene="At3g01530/F4P13_8"
/codon_start=1
/product="putative transcription factor"
/protein_id="BAC42719.1"
/db_xref="GI:26451234"
/translators="METTMKKGRVKATITTSQEEEGTVRGKPMTEEDFILFYIILN
HGEIWMNSVAKASGLKRTGSKRLWNLPRDVRGNITTEEOILLIOLHAKLGNRW
SKIAHLPGRTDNEIKNFWRTKIQHMKVSENMMHCHCSGNSQSGMTTQSSSK
AIDTAESFSQAKTTFNVVEQQSNENYVWEDLWVHLNGDHHVI"

ORIGIN
Query Match 23.7%; Score 204.8; DB 8; Length 894;
Best Local Similarity 76.5%; Pred. No. 5.7e-34;
Matches 251; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 60 AAGATCTCGAAGTGAAGAAAGGGCTTGGACAATGGGAAGAGCTTGATCTTGATGAAGT 119
Db 170 AAGAAGGAACAGTGAAGAAAGGACCATGCACTATGGAAGAAGATTTCATCCTCTTTAAT 229

Qy 120 ATATTGCAATCATCGGGAAGGTCTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAC 179
Db 230 ACATCCCTTAATCATGCTGGTGAAGTCTTTGGAACTCTGTGCGCAAAAGCCTCTGGTCTAAAC 289

Qy 180 GTAAACGGAAGAGTTGCCGGCTTAAGTGGCTTAAATTAACCTCCGTCCTGATCTTAGAAGAG 239
Db 290 GTACTGGAAGAAGTTGTCCGCTCCGGTGGCTGAATCTATCCGACCAAGATGTGCGGCGAG 349

Qy 240 GGAATATTACACCGAGGAACAATTTTGGATTATGGAGCTCCAGCAAGTGGGGAACA 299
Db 350 GGAACATAACCGAAGAAGAACAGCTTTTGATCAATTCAGCTTCAAGCTTGGAACA 409

Qy 300 GGTGGTCCAAATTCGCAAGCATCTACTGGAAGGCTGATAATGAGTCAAGAACTATT 359
Db 410 GGTGGTCCGAATTCGGAAGCATCTTCGGGAAGACGGACACGAGATAAAGAACTTCT 469

Qy 360 GGAGGACAAGGATCCAGGACGATCAAC 387
Db 470 GGAGGACAAGATTCAGAGACATGAA 497

RESULT 14
AB058642
LOCUS
DEFINITION
AB058642
VERSION
AB058642.1 GI:13537529
KEYWORDS
SOURCE
ORGANISM
Lilium hybrid division I
Lilium hybrid division I
Lilium hybrid division I
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
Lilium; Lilium hybrid cultivars.
1 (sites)
Nakatsuka, A., Izumi, Y. and Yamagishi, M.
Isolation and characterization of the genes related to anthocyanin
biosynthesis in Asiatic hybrid lily
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 872)
AUTHORS
Nakatsuka, A., Izumi, Y. and Yamagishi, M.
TITLE
Direct Submission
JOURNAL
Submitted (24-MAR-2001) Akira Nakatsuka, Shimane University,
Faculty of Life and Environmental Sciences; Nisikawatsu 1060,
Matsue, Shimane 690-8504, Japan
(E-mail: nakira@life.shimane-u.ac.jp, Tel: 81-852-32-6502)
Location/Qualifiers
FEATURES
1. .872
source
/organism="Lilium hybrid division I"
/mol_type="mRNA"
/cultivar="Montreux"

/db_xref="taxon:156532"
/tissue_type="tepai"
1. .872
/gene="LhMyb"
80. .634
/gene="LhMyb"
/codon_start=1
/protein_id="BAB40790.1"
/db_xref="GI:13537530"
/translators="MDRKVTNEDVEVRKGPWTMEEDLILINYIANHGDGVWNTLARS
AGLKRTGSKRLWNLPRDVRGNITTEEOILLIOLHAKLGNRW
SKIAHLPGRTDNEIKNFWRTKIQHMKVSENMMHCHCSGNSQSGMTTQSSSK
AIDTAESFSQAKTTFNVVEQQSNENYVWEDLWVHLNGDHHVI"

ORIGIN
Query Match 22.5%; Score 194.6; DB 8; Length 872;
Best Local Similarity 69.1%; Pred. No. 9.2e-32;
Matches 266; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 52 GAGCTCTCAAGATCTCTGAAGTGAAGAAAGGGCTTGGACAATGGGAAGAGCTTGATCTT 111
Db 97 GACTAATGAAGATGTTGAGGTGAGGAAGGACCATGCGACGAGTCTGGAACACGCTGCTCGATCGGCGG 156

Qy 112 GATGAACATATTTGCAATCATGGGAAGGTGTTGAACTCTTTGGCCAAAGCTGCTGG 171
Db 157 CATCAACTACATAGCCCAACCATGCGACGAGTCTGGAACACGCTGCTCGATCGGCGG 216

Qy 172 TCTCAAAAGCTTAACGGAAGAGTTGCCGGCTTAAGTGGCTTAAATTAACCTCCGTCCTGATGT 231
Db 217 ACTGAAGAGAGCTTGGAAAGAGCTGCGAGCTTCGTTGGCTGAATTAACCTGAGGCTGATGT 276

Qy 232 TAGAAGAGGAATATTATACCCGAGGAACAACCTTTTGAATTTATGAGTCTCCACGCAAGTG 291
Db 277 TCGTGGGGTAAACATACACACGAGAGGAGCAGCTTCTGATCATGGAACCTTCACGACAGATG 336

Qy 292 GGGAAACAGCTGCTCCAAATTCGCAAGCATCTACCTGGAGGAGCTATATGAGATGAGATCA 351
Db 337 GGSTAAACAGATGCTGCAAAATTCGCAAGGAGCTCCAGGACGAAACAGATACGAAATCAA 396

Qy 352 GAACATTTGGAGGACAGGATCCAGAGCAGCATCAAGCAAGCTGAGAATTTTCAGCAACA 411
Db 397 GAACCTTCTGGAGGACGAGTCCAGAGAATCCAGAAATGGCGAATCATCTGAGGGCCA 456

Qy 412 GAGTAGTAATTAATTTCTGAGATAAT 436
Db 457 AATCTCCATGCTTTGGATGAAACT 481

RESULT 15
AF485900
LOCUS
DEFINITION
Dendrobium sp. XMW-2002-9 MYB9 (MYB9) mRNA, complete cds.
ACCESSION
AF485900
VERSION
AF485900.1 GI:28628962
KEYWORDS
SOURCE
ORGANISM
Dendrobium sp. XMW-2002-9
Dendrobium sp. XMW-2002-9
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
Epidendroideae; higher Epidendroideae; Dendrobieae; Dendrobieae;
Dendrobium.
1 (bases 1 to 649)
AUTHORS
Wu, X.-M., Lim, S.-H. and Yang, W.-C.
TITLE
Characterization, expression and phylogenetic study of R2R3-MYB
genes in orchid
JOURNAL
Plant Mol. Biol. 51 (6), 959-972 (2003)
MEDLINE
22661209
PUBMED
12777054
REFERENCE
2 (bases 1 to 649)
AUTHORS
Wu, X.-M., Lim, S.-H. and Yang, W.-C.
TITLE
Direct Submission
JOURNAL
Submitted (21-FEB-2002) Laboratory of Plant Reproduction & Cotton
Biotechnology, Institute of Molecular Agrobiology, 1 Research Link
```

FEATURES

S117604, Singapore

Location/Qualifiers

1..649

/organism="Dendrobium sp. XMW-2002-9"

/mol_type="mRNA"

/db_xref="taxon:217455"

1..649

/gene="MYB9"

21..476

/gene="MYB9"

/note="MYB"

/codon_start=1

/product="MYB9"

/protein_id="AAO49418.1"

/db_xref="GI:28628963"

/translation="MEEDGILINVIANHGEVWNNLAKSAGLKRTGKSCRLLRWNLRLR

PDRVRCNLTPEQLLIMELHARWGNRWSKIARQLPGRDNEIKYWRTRIQKKVNGD

AVDYSSQMIIDEACTSHTSYSSIEEGMAQQONYPVHPNTNPEGYRLQTRVN"

ORIGIN

Query Match

22.0%; Score 190; DB 8; Length 649;

Best Local Similarity 75.8%; Pred No. 9.5e-31;

Matches 235; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY

87

GGCAATGGAAGAAGACTTGATCTTGTGATGAACATATATTGCAAAATCATGGGGAAGGTGTTT

146

Db

16

GGACTATGGAAGAGGACGGGATCCTCATAACTATATAGCCAAACCATGGAGAGGGAGTAT

75

QY

147

GGAACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGGAAAGAGTTGCCGGCTAAGGT

206

Db

76

GGAAATAATCTTGCCAAATCTGCAGGTTTGAAGAGGACGGGGAAGAGCTGCCGGCTTCGAT

135

QY

207

GGCTAAATTTACCTCCTCGTCTGATGTTAGAAGAGGGAATATTACACCCGAGGAACAACTTT

266

Db

136

GGCTAAACTATCTGAGGCCAGACGTCGAGAGGGGGAACATACACCGAGAGGACGCTGC

195

QY

267

TGATTATGGAGCTCCACGCAAAAGTGGGGAACAGGTGGTCCAAAATTGCCAAGCATCTAC

326

Db

196

TCATCATGGAGTTGCATGCCAGTGGGGAACAGGTGGTCAAAAATTGCGAGGCAACTTC

255

QY

327

CTGGAAGGACTGATTAATGAGATCAAGAACTATTGGAGGACAAAGGATCCAGAGGACATCA

386

Db

256

CAGGAAGGACTGATAATGAGATCAAGAACTACTGGAGGACTAGAAATTCAAAAGAAGTTA

315

QY

387

AGCAAGCTGA

396

Db

316

AGATGGCGA

325

Search completed: September 28, 2005, 10:37:59
Job time : 4121 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 06:36:16 ; Search time 585 Seconds
(without alignments)
8732.883 Million cell updates/sec

Title: US-10-659-869A-35
Perfect score: 863
Sequence: 1 gcacgagctctatcacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861.8	99.9	863	8	ACD07386 Soybean M
2	861.8	99.9	863	12	ADJ77736
3	800.4	92.7	968	12	ADI43074 Plant tra
4	800.4	92.7	968	12	ADO03218 Soybean o
5	610.2	70.7	988	8	ACD07383
6	610.2	70.7	988	12	ADJ77730 cDNA enco
7	423	49.0	969	12	ADI43075 Plant tra
8	423	49.0	969	12	ADO03219 Soybean o
9	400.2	46.4	988	12	ADI43073 Plant tra
10	400.2	46.4	988	12	ADO03217 Soybean o
11	394.4	45.7	805	8	ACD07387
12	394.4	45.7	805	12	ADJ77738
13	342.6	39.7	910	8	ACD07385
14	342.6	39.7	910	12	ADJ77734
15	303.8	35.2	530	8	ACD07384
16	303.8	35.2	530	12	ADJ77732 cDNA enco
17	298.8	34.6	392	12	ADO06964
18	287.4	33.3	557	8	ACD07382
19	287.4	33.3	557	12	ADJ77728
20	273.6	31.7	681	10	ADD17714 DNA (seqi

21	273.6	31.7	681	10	ADK56061
22	265.4	30.8	681	12	ADJ10801
23	250.8	29.1	462	10	ADD17692
24	250.8	29.1	462	10	ADK59524
25	248.4	28.8	862	3	AAC57336
26	239.4	27.7	910	3	AAC40587
27	236.2	27.4	908	3	AAC48506
28	234.6	27.2	731	12	ADI43518
29	234.6	27.2	731	12	ADO03344
30	234.6	27.2	731	12	ADO01760
31	223	25.8	907	3	AAC57337
32	215	24.9	1033	9	ACD98402
33	215	24.9	1033	10	ADD55727
34	215	24.9	1033	10	ADD30419
35	215	24.9	1033	12	ADI43506
36	215	24.9	1033	12	ADO03332
37	209	24.2	235	12	ADO05985
38	208	24.1	621	3	AAC41714
39	208	24.1	1127	3	AAC37953
40	206.4	23.9	1011	10	ADD30197
41	206.4	23.9	1011	10	ADE37062
42	206.4	23.9	1011	10	ADE37082
43	206.4	23.9	1011	10	ADE31530
44	206.4	23.9	1011	12	ADI41794
45	206.4	23.9	1011	12	ADI61430

ALIGNMENTS

RESULT 1
ACD07386
ID ACD07386 standard; cDNA; 863 BP.
XX
AC ACD07386;
XX
AC ACD07386;
DT 07-AUG-2003 (first entry)
XX
DE Soybean Myb-related transcription factor #5 cDNA.
XX
KW Soyben; ss; Myb-related transcription factor; crop improvement;
KW starch composition; protein composition; plant; gene; grain quality;
KW disease resistance; pest resistance; oil composition.
XX
OS Glycine max.
XX
PN US2003024007-A1.
XX
PD 30-JAN-2003.
XX
PF 14-DEC-2001; 2001US-00021811.
XX
PR 02-DEC-1998; 98US-0110609P.
PR 01-DEC-1999; 99US-00452244.
XX
(CAHO/) CAHOON R E.
(ODEL/) ODELL J T.
XX
PI Cahoon RE, Odell JT;
XX
XX WPI; 2003-456317/43.
DR P-PSDB; ABO01759.
XX
XX New isolated polynucleotide encoding a Myb-related transcription factor,
PT useful in plant molecular biology, in particular for the improvement of
PT crop plants for a variety of traits, including disease and pest
PT resistance.
XX
PS Claim 2; Page 36; 61pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a Myb-
CC related transcription factor polypeptide. The methods and compositions of
CC the present invention are useful in plant molecular biology, in

particular for the improvement of crop plants for a variety of traits, including disease and pest resistance and grain quality improvements such as oil, starch or protein composition. The present sequence represents cDNA encoding a soybean Myb-related transcription factor

XX
SQ Sequence 863 BP; 306 A; 149 C; 180 G; 225 T; 0 U; 3 Other;

```
Query Match      99.9%; Score 861.8; DB 8; Length 863;
Best Local Similarity 100.0%; Pred. No. 3.9e-163;
Matches 863; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy 1 GCACGAGCTCTATCACACACACAGTCAATGGATAAAAAACAACAGTGTAAAGACGCTCTCA 60

Dβ 1 GCACGAGCTCTATCACACACACAGTCAATGGATAAAAAACAACAGTGTAAAGACGCTCTCA 60

61 AGATCCTGAGTGAGAGAAAGGGCCCTGGACAATGGAGCAAGACCTTGCTCTGTGAACTA 120

D_b 61 AGATCCTCGAAGTGCAGAAAGGGCCWTTGGA CAATGGGAAGAACA CTTTGATCTTGATGAACCTA 120

121 TATTGAAATCATGGGGAAGGTGTGTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAACG 180

D6 121 TATTGCAAAATCATGGGGAAGGTGTTTGGAACTCTTTGGCAAAGCTGCTGTTCTCAAACG 180

181 TAA CGG AAG AGT TGC CGG CTA AGT TGG CTA AAT TAC TCC GTC GAT GTT A GA AGG 240

Db 181 TAACGGAAAGAGTTGCCGGCTAAGTGGCTAAATTACCTCCGTCCTGATGTTAGAGAGG 240

QV 241 GAATATTACACCCGAGGAACAACCTTTTGTATTGGAGCTCCACGCAAAAGTGGGGAACAG 300

Db 241 GAATATTACCCGAGGAACAACCTTTTGTATGGAGCTCCACGAAAGTGGGAAACAG 300

QV 301 GTGGTCCAAAATTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTG 360

Db 301 GTGGTCCAAATTTGCCAAGCATCTTACCTGGAAGGACTGATAATGACATCAAGAACTATTG 360

QV 361 GAGGACAAGGATCCAGAGGCACATCAAGCAAGCTGAGAACTTTCACCAACACAGAGTAGTAA 420

Db 361 GAGGACAAGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTGAGCAACAGATAGTAA 420

QY 421 TAATTTCTGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCC 480

Db 421 TAAATTCGAGATAAATCATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCC 480

Qy 481 CATGGAGATGTATTCTCCACCCCTGTTATCAAGGAATGTTAGGCCATTTTCAACTCAGTT 540

Db 481 CATGGAGATGTTATCTCCACCCCTGTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTT 540

QY 541 CCTACAAATTAATCCTGATCAATCCAGTTGTTGATCCAAATGACAACAACACATTAAC 600

Db 541 CCCTACAATTATCTCGATCAATCCAGTTGTTGTACCAATGACAACAACACATTAACTA 600

QY 601 TTGGAGCATGGAGGATACCTGGTCAATGCAATTACTGACGGTGATTAATAATTATCAAG 660

Db 601 TTGGAGCATGGAGGATAGTGGTCAATGCAATTACTGAACGGTGATTAATAATTATCAAG 660

QY 661 ATAAACCTAAGTTTGAAGTTCATAGGCTGGAATGTCCTYTGGAATTAACATATTAT 720

Db 661 ATAAAACCTAAGTTTGTAGTTCCTAAGCTGGAATGCTCTTGTGGATTAAAAACATATTAT 720

Qy 721 TGGGTTTGTATATAAGTAGTTGGATGTTTGGTTTTGCGTACCATTTATTAGCTATGTGC 780

Db 721 TGGGTTTGTTATATAAGTAGTTGGATGTTTGGTTTGGTACCATATTAGCTATGTC 780

QY 781 TGTAATATATACGAGATYTTATATTAACTATATCTGCATGCTTTATATATATAAAAAA 840

Db 781 TGTAATATATACGAGATVTTATATATAAACTATATCTGCATGCTTTATATATAAAAAAAA 840

Qy 841 AAAAAAAAAAAAAAAAAAAAAA 863

Db 841 AAAAAAAAAAAAAAAAAAAAAA 863

RESULT 2

QY

181 TAACGGAGAGATTGCCCCGCTAAGGTGGCTAAATATACCTCCGTCCTGATGTATAGAAAGAGG 249


```
Db 181 TAACGGAAGAGTTGCGGGCTAAGGTGCTAAATTTACCTCGTCTGTATGTAAGAGG 240
Qy 241 GAATATTACACCCGAGGAAACAATTTTGTATTTGAGCTCCACGAAAGTGGGAAACAG 300
Db 241 GAATATTACACCCGAGGAAACAATTTTGTATTTGAGCTCCACGAAAGTGGGAAACAG 300
Qy 301 GTGTTCAAATTTGCCAAGCATCTACCTGGAAGCTGTGTAATGAGATCAAGAACTATTG 360
Db 301 GTGTTCAAATTTGCCAAGCATCTACCTGGAAGCTGTGTAATGAGATCAAGAACTATTG 360
Qy 361 GAGGACAAGATCAGAGACACATCAAGCAAGCTGAGCACTTTTCAGCAACAGAGTAGTAA 420
Db 361 GAGGACAAGATCAGAGACACATCAAGCAAGCTGAGCACTTTTCAGCAACAGAGTAGTAA 420
Qy 421 TAATTTCTGAGATAAATGATCACCAGCTAGCACTAGCACTGTTTCCACCATTGGCTGAGCC 480
Db 421 TAATTTCTGAGATAAATGATCACCAGCTAGCACTAGCACTGTTTCCACCATTGGCTGAGCC 480
Qy 481 CATGGAGATGTAATCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTT 540
Db 481 CATGGAGATGTAATCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTT 540
Qy 541 CCCTACAAATTAATCCTGATCAATCCAGTTGTTGTACCAATGACACACACATTAACCTA 600
Db 541 CCCTACAAATTAATCCTGATCAATCCAGTTGTTGTACCAATGACACACACATTAACCTA 600
Qy 601 TTGAGCATGAGGATAGCTGGTCAATGCAATTAATGCAAGCTGATTAATTAATTAATCAAG 660
Db 601 TTGAGCATGAGGATAGCTGGTCAATGCAATTAATGCAAGCTGATTAATTAATTAATCAAG 660
Qy 661 ATAAACCTAAGTTTGAAGTTCCATAAGCTGGAATGTTCTGTGATTAATAAATAATAT 720
Db 661 ATAAACCTAAGTTTGAAGTTCCATAAGCTGGAATGTTCTGTGATTAATAAATAATAT 720
Qy 721 TGGGTTTGTATATAGTAGTTGGATGTTTGGTTTGGCTGACCATTAATGCTATGTC 780
Db 721 TGGGTTTGTATATAGTAGTTGGATGTTTGGTTTGGCTGACCATTAATGCTATGTC 780
Qy 781 TGTAAATATACGAGATTTATATTAACCTATATCTGCATGCTTTATATATAAATAAATA 840
Db 781 TGTAAATATACGAGATTTATATTAACCTATATCTGCATGCTTTATATATAAATAAATA 840
Qy 841 AAAAAAAAAAAAAAAAAAAAAA 863
Db 841 AAAAAAAAAAAAAAAAAAAAAA 863

RESULT 3
ADI43074
ID 4 ADI43074 standard; DNA; 968 BP.
AC ADI43074;
XX
XX 22-APR-2004 (first entry)
DT
DE
DE Plant transcription factor polynucleotide #1009.
XX
XX transgenic; plant; enhanced tolerance to abiotic stress;
KW glyphosate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
XX
XX Glycine max.
XX
XX US2004019927-A1.
XX
XX 29-JAN-2004.
PD
```

```
XX 25-FEB-2003; 2003US-00374780.
XX 18-APR-2001; 2001US-00837944.
XX (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX
XX WPI; 2004-132245/13.
XX
XX New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX
XX Claim 1; SEQ ID NO 1537; 435pp; English.
XX
XX The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produce a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; sugar sensing; early or late flowering;
CC sensitivity; disease resistance; glycosylated tolerance; hormone
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanins, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant
CC transcription factor, and an orthologue of Arabidopsis thaliana
CC transcription factors isolated in the invention, that can be used in the
CC creation of a transgenic plant with altered traits.
XX
XX Sequence 968 BP; 308 A; 197 C; 195 G; 268 T; 0 U; 0 Other;
SQ
XX
XX Query Match 92.7%; Score 800.4; DB 12; Length 968;
XX Best Local Similarity 97.4%; Pred. No. 7.7e-151;
XX Matches 810; Conservative 3; Mismatches 19; Indels 0; Gaps 0;
Qy 8 CTCTATCACACACACAGTCAATGGATATAAAAAACACAGGTGTAGAGCTTCAAGATCCT 67
Db 51 CTCTATCACACACACAGTCAATGGATATAAAAAACACAGGTGTAGAGATCCT 110
Qy 68 GAAGTGAAGAAAGGGCTTGGACATGGAAGAGAGCTTGTATGTAAGTATATTGCA 127
Db 111 GAAGTGAAGAAAGGGCTTGGACATGGAAGAGAGAGCTTGTATGTAAGTATATTGCA 170
Qy 128 AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAACACGTAACGGA 187
Db 171 AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAACACGTAACGGA 230
```


CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed morphology, increased cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenyl lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence encodes an
 CC orthologue of a thalecress transcription factor isolated from Soybean.

SQ Sequence 968 BP; 308 A; 197 C; 195 G; 268 T; 0 U; 0 Other;

Query Match 92.7%; Score 800.4; DB 12; Length 968;
 Best Local Similarity 97.4%; Pred. No. 7.7e-151;
 Matches 810; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

Qy 8 CTCATATCACACACAGTCAATGGATAAAAAACACAGTGTAAAGAGTCTCAAGATCCT 67
 Db 51 CTCATATCACACACAGTCAATGGATAAAAAACACAGTGTAAAGAGTCTCAAGATCCT 110
 Qy 68 GAAGTGAGAAAGGCGCTTGACAAATGGAAGACATTTGATGTAAGTAACTATATTGCA 127
 Db 111 GAAGTGAGAAAGGCGCTTGACAAATGGAAGACATTTGATGTAAGTAACTATATTGCA 170
 Qy 128 AATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGCTCAACAGTAAACGGA 187
 Db 171 AATCATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGCTCAACAGTAAACGGA 230
 Qy 188 AAGAGTTCGGGCTAAAGTGGCTAAATTAATCTCGCTCTGATGTAGAGGGGAATATT 247
 Db 231 AAGAGTTCGGGCTAAAGTGGCTAAATTAATCTCGCTCTGATGTAGAGGGGAATATT 290
 Qy 248 ACACCCGAGGAACTATTTGATTTATGAGCTCCACGAAAGTGGGAAACAGTGTGTC 307
 Db 291 ACACCCGAGGAACTATTTGATTTATGAGCTTCACGAAAGTGGGAAACAGTGTGTC 350
 Qy 308 AAAATTGCAAGCATCTACCTGGGAAGGACTGATATGAGATCAAGAACTATTGAGGACA 367
 Db 351 AAAATTGCAAGCATCTACCTGGGAAGGACTGATATGAGATCAAGAACTATTGAGGACA 410
 Qy 368 AGGATCCAGAGCAATCAAGCAAGCTGAGAACTTTTTCAGCAACAGAGTAGTAATTTCT 427
 Db 411 AGGATCCAGAGCAATCAAGCAAGCTGAGAACTTTTTCAGCAACAGAGTAGTAATTTCT 470
 Qy 428 GAGATAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 487
 Db 471 GAGATAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 530
 Qy 488 ATGATTTCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTTACA 547
 Db 531 ACCTATTTCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTTACA 590
 Qy 548 ATTAATCTCTGATCAATCCAGTTGTTGTACCAATGACAAACAACTTAATCTATTGAGC 607

Db 591 ATTAATCTCTGATCAATCCAGTTGTTGTACCAATGACAAACAACTTAATCTATTGAGC 650
 Qy 608 ATGGAGGATAGCTGGTCAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 667
 Db 651 ATGGAGGATAGCTGGTCAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 710
 Qy 668 CTAAGTTCGAAGTTCCTAAGGCTGGAATGCTCTCTGGAATTAATTAATTAATTAATTA 727
 Db 711 CTAAGTTCGAAGTTCCTAAGGCTGGAATGCTCTCTGGAATTAATTAATTAATTAATTA 770
 Qy 728 GTTTATATTAAGTAGTTGGATGTTGTTGTTGCTTACCATTTATTAATTAATTAATTA 787
 Db 771 GTTTATATTAAGTAGTTGGATGTTGTTGTTGCTTACCATTTATTAATTAATTAATTA 830
 Qy 788 TATACGAGATTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839
 Db 831 TATACGAGATTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 882

RESULT 5
 ACD07383

ID ACD07383 standard; cDNA; 988 BP.

XX ACD07383;

XX 07-AUG-2003 (first entry)

DE Soybean Myb-related transcription factor #2 cDNA.

XX Soyben; ss; Myb-related transcription factor; crop improvement;
 KW starch composition; protein composition; plant; gene; grain quality;
 KW disease resistance; pest resistance; oil composition.

XX Glycine max.

PN US2003024007-A1.

XX 30-JAN-2003.

XX 14-DEC-2001; 2001US-00021811.

XX 02-DEC-1998; 98US-0110609P.

XX 01-DEC-1999; 99US-00452244.

XX (CAHO//) CAHOON R E.

PA (ODEL//) ODELL J T.

XX Cahoon RE, Odell JT;

XX WPI; 2003-456317/43.

DR P-PSDB; ABO01756.

XX New isolated polynucleotide encoding a Myb-related transcription factor,
 PT useful in plant molecular biology, in particular for the improvement of
 PT crop plants for a variety of traits, including disease and pest
 PT resistance.

XX Claim 2; Page 32-33; 61pp; English.

XX The invention relates to an isolated polynucleotide encoding a Myb-
 CC related transcription factor polypeptide. The methods and compositions of
 CC the present invention are useful in plant molecular biology, in
 CC particular for the improvement of crop plants for a variety of traits,
 CC including disease and pest resistance and grain quality improvements such
 CC as oil, starch or protein composition. The present sequence represents
 CC cDNA encoding a soybean Myb-related transcription factor

SQ Sequence 988 BP; 339 A; 195 C; 182 G; 272 T; 0 U; 0 Other;

Query Match 70.7%; Score 610.2; DB 8; Length 988;

Best Local Similarity 92.0%; Pred. No. 9.2e-113;

Matches 724; Conservative 2; Mismatches 40; Indels 21; Gaps 7;

```
QY      8  CTCTATCA CACACACAAAGTCAATGATATAAAAAACAAAGTGTAAAGACGCTCTCAAGATCCT 67
Db      40  CTCTATCA CACACACAAACAAATGATATAAAAAACAACTGTGCAACACGCTCTCAAGATCCT 99

QY      68  GAACTGAGAAAGGCGCTTGACATGGAAGAGACTTGGATCTTGATGAACATATATGCA 127
Db      100  GAACTGAGAAAGGCGCTTGACATGGAAGAGACTTGGATCTTGATGAACATATATGCA 159

QY      128  AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAAACGCTAAACGGA 187
Db      160  AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAAACGCTAAACGGA 219

QY      188  AAGAGTTCGCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 247
Db      220  AAGAGTTCGCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 279

QY      248  ACACCCGAGGAAACAACTTTTGGATGAGCTCCACGCAAAAGTGGGGAACAGGTGGTCC 307
Db      280  ACACCCGAGGAAACAACTTTTGGATGAGCTTCACGCAAAAGTGGGGAACAGGTGGTCC 339

QY      308  AAAATTGCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACA 367
Db      340  AAAATTGCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACC 399

QY      368  AGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATATCT 427
Db      400  AGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAAATCAGCAATAACTCT 459

QY      428  GAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACATGGCTGAGCCCATGGAG 487
Db      460  GAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACATGGCTGAAACCCATGGAG 519

QY      488  ATGTATTCTCCACCGCTGTTATCAAGGAATGTTAGAGCCATT--TTCAACTCAGTTCCT 544
Db      520  ACCTATTCTCCACCGCTTATCAAGGAATGTTAGAGCCATTTCTCAATTCAGTTCCCC 579

QY      545  ACAATTAATCTCGATCAATCCAGTTGTTGTACCAATGACAACAA--CATTAACATAT 601
Db      580  ACAATTAATCTCGATCAATCCAGTTGTTGTACCAATGACAACAAACAGCATTAACATAT 639

QY      602  TGGAGCATGGAGGATAGCTGTCATGCAATTAATCTGAACGGTGATTAATAT--TAT 656
Db      640  TGGAGCATGGAGGATATCTGGTCAATGCAATGCAATGCAATGCAATGCAATGCAAT 699

QY      657  CAAGATAAAACCTAAGTTTGTG--AAGTTTCCAT--AAGCTGGAATGCTTGGATTAAA 711
Db      700  CAAGATAAAACCTAATTTCTGTATTAAGTTCCATAAAACACTGGAATGCTCTGGCTTAAA 759

QY      712  ACA---TATTATGGGTTTGTATTATATAAGTAGTTGG--ATGTTTGGTTTTCGCPACCAT 766
Db      760  ACATATTATTATTAGGTTTGTATTATATAAGTAGTTGGATATGTTTGGTTTTCGCPACCAT 819

QY      767  TATTAGC 773
Db      820  TATTAGC 826
```

RESULT 6

```
ADJ77730
ID   ADJ77730 standard; cdNA; 988 BP.
XX
AC   ADJ77730;
XX
CC   ADJ77730;
XX
DT   20-MAY-2004 (first entry)
XX
DE   cdNA encoding soybean Myb-related transcription factor #2.
XX
KW   Plant; Myb-related transcription factor; disease resistance; soybean;
KW   gene; ss.
XX
OS   Glycine max.
XX
PN   US2004040057-A1.

Query Match      70.7%; Score 610.2; DB 12; Length 988;
Best Local Similarity 92.0%; Pred. No. 9.2e-113;
Matches 724; Conservative 2; Mismatches 40; Indels 21; Gaps 7;

QY      8  CTCTATCA CACACACAAAGTCAATGATATAAAAAACAAAGTGTAAAGACGCTCTCAAGATCCT 67
Db      40  CTCTATCA CACACACAAACAAATGATATAAAAAACAACTGTGCAACACGCTCTCAAGATCCT 99

QY      68  GAACTGAGAAAGGCGCTTGGAACAATGGAAGAAGACTTTGATCTTTGATGAACATATATGCA 127
Db      100  GAACTGAGAAAGGCGCTTGGAACAATGGAAGAAGACTTTGATCTTTGATCAACTATATGCA 159

QY      128  AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAAACGCTAAACGGA 187
Db      160  AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAAACGCTAAACGGA 219

QY      188  AAGAGTTCGCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 247
Db      220  AAGAGTTCGCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 279

QY      248  ACACCCGAGGAAACAACTTTTGGATGAGCTCCACGCAAAAGTGGGGAACAGGTGGTCC 307
Db      280  ACACCCGAGGAAACAACTTTTGGATGAGCTTCACGCAAAAGTGGGGAACAGGTGGTCC 339

QY      308  AAAATTGCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACA 367
Db      340  AAAATTGCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACC 399

QY      368  AGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATATCT 427
Db      400  AGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAAATCAGCAATAACTCT 459
```

```
XX      26-FEB-2004.
PD
XX
XX
PF      11-SEP-2003; 2003US-00659869.
XX
XX      02-DEC-1998; 98US-0110609P.
PR      01-DEC-1999; 99US-00452244.
PR      14-DEC-2001; 2001US-00021811.
XX
XX      (CAHO/) CAHOON R E.
PA      (FANG/) FANG Y.
PA      (ODEL/) ODELL J T.
PA      (WENG/) WENG Z.
XX
PI      Cahoon RE, Fang Y, Odell JT, Weng Z;
XX
XX      WPI; 2004-238121/22.
DR      P-PSDB; ADJ77731.
DR
XX
```

Plant Myb transcription factor homologs and their polynucleotide sequences useful in plant genomic research e.g. identifying nucleotide fragments that encode a Myb-related transcription factor polypeptide.

Claim 2; SEQ ID NO 29; 62pp; English.

The present invention relates to the isolation of polynucleotide sequences encoding plant Myb-related transcription factors and the encoding polypeptide sequences. Also disclosed is a chimeric gene encoding all or part of a Myb-related transcription factor linked to suitable regulatory sequences, and a host cell transformed with the chimeric gene and able to express the chimeric gene resulting in the production of altered levels of the Myb-related transcription factor in the transformed cell. The polynucleotide sequences of the invention may be used to probe a cDNA or genomic library to identify nucleic acid fragments encoding a Myb-related transcription factor polypeptide and also for research purposes including generating plants with increased resistance to disease, etc. The present sequence encodes a plant Myb-related transcription factor.

Sequence 988 BP; 339 A; 195 C; 182 G; 272 T; 0 U; 0 Other;

```
QY 428 GAGATAAATGATACCAAGCTAGCAGCTAGCCATGTTTCCACATGGCTGAGCCCATGGAG 487
D5 460 GAGATAAATGATACCAAGCTAGCAGCTAGCCATGTTTCTACCATGGCTGAACCCCATGGAG 519
QY 488 ATGATTCTCCACCTGTTATCAAGGAATGTTAGAGCCATT--TTCAACTCAGTTCCT 544
D5 520 ACCTATTCTCACCTTTTATCAAGGAATGTTAGAGCCATTTCCTCAATTCAGTTCCTCC 579
QY 545 ACAATTAATCTGATCAATCCAGTTGTTGTACCAATGACAACAACAA--CATTAACCTAT 601
D5 580 ACAATTAATCTGATCAATCCAGTTGTTGTACCAATGACAACAACAAACAGCATTAACCTAT 639
QY 602 TGGAGCATGGAGGATAGCTGTCATGCAATTAAGCTGATTAATAAT--TAT 656
D5 640 TGGAGCATGGAGGATATCTGTCATGCAATGAGTACTGTAACGGGGATTAATAATGATATAT 699
QY 657 CAAGATAAACTAAGTTTTC--AAGTTCCAT--AAGCGTGAATGCTCTGTGATTAA 711
D5 700 CAAGATAAACTAATCTCTGTATTAAGTTCCATAAAGCTGGAATGCTCTGCTTAA 759
QY 712 ACA---TATTATGGTTTGTATATATAAGTAGTTGG--ATGTTTGGTTTTCGTACCAT 766
D5 760 ACATATTATTATAGTTTGTATATAAGTAGTTGGATATGTTTGGTTTTCGTACCAT 819
QY 767 TATTAGC 773
D5 820 TATTAGC 826

RESULT 7
AD143075
ID AD143075 standard; DNA; 969 BP.
XX AC AD143075;
XX DT 22-APR-2004 (first entry)
XX DE Plant transcription factor polynucleotide #1010.
XX KW transgenic plant; enhanced tolerance to abiotic stress;
KW glycosphatic tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; gene; ds.
XX OS Glycine max.
XX PA US2004019927-A1.
XX FN 29-JAN-2004.
XX PD 25-FEB-2003; 2003US-00374780.
XX PF 18-APR-2001; 2001US-00837944.
XX PR (SHER/) SHERMAN B K.
XX PA (RIEC/) RIECHMANN J L.
XX PA (JIAN/) JIANG C.
XX PA (HEAR/) HEARD J E.
XX PA (HAAR/) HAAKE V.
XX PA (CREE/) CREELMAN R A.
XX PA (RATC/) RATCLIFFE O.
XX PA (ADAM/) ADAM L J.
XX PA (REUB/) REUBER T L.
XX PA (KEDD/) KEDDIE J.
XX PA (BROU/) BROUN P E.
XX PA (PILG/) PILGRIM M L.
XX PA (DUBE/) DUBELL A N.
XX PA (PINE/) PINEDA O.
```

```
PA (YUGG/) YU G.
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
XX New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX Claim 1; SEQ ID NO 1538; 435pp; English.
XX The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produce a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glycosphatic tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanin; increase in plant
CC anthocyanin; or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant
CC transcription factor, and an orthologue of Arabidopsis thaliana
CC transcription factors isolated in the invention, that can be used in the
CC creation of a transgenic plant with altered traits.
XX SQ Sequence 969 BP; 311 A; 188 C; 200 G; 270 T; 0 U; 0 Other;
Query Match 49.0%; Score 423; DB 12; Length 969;
Best Local Similarity 81.5%; Pred. No. 2.7e-75;
Matches 573; Conservative 2; Mismatches 92; Indels 36; Gaps 6;
QY 81 GGCCTTGGACAATGAAGAAGACTTGATCTTGATGAACATATATTGCAAAATCATCGGGAG 140
D5 44 GGGCAGGACAATGAAGAAGACTTAATCTTGATCACCCTATATTGGCAATCACGGGAG 103
QY 141 GTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACGTAACGGAAGAGTTGCCGGC 200
D5 104 GGGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACGTAACGGAAGAGTTGCCGGC 163
QY 201 TAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAAGAGGAATATTACACCCGAGGAAC 260
D5 164 TTCGGTAGCTAAACTACCTACCTAGCTCCTGATGTTAGAAGAGGAATATTACACCCGAGGAAC 223
QY 261 AACTTTTGATATGGAGCTCCAGCAAGTGGGAAACAGGTGTCCTCAAAATTCGCAAGC 320
D5 224 AGCTTTTGATCATGGAACTTTCATGCAAAAGTGGGCAACAGGTGTCCTCAAAATTCGCAAGC 283
QY 321 ATCTACCTGGAAGACTGATAATGAGATCAAGAACTATTGGAGCAACAGGATCCAGAAGC 380
D5 284 ATCTACCCGGAAGACTGATAATGAGATTAAGAACTACTGGAGCAACAGGATCCAGAAGC 343
QY 381 ACATCAAGCAAGCT--GAGAACCTTTTCAGCAACAGAGTAGTAATAATTCTGAGATAAATG 437
D5 344 ACCTCAAGCAAGCTTCCAGCAGCTTCCAGCAACAGAGTAGTAATAATTCTGAGATAAATG 403
QY 438 ATCAAGCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGAGATGATTCTC 497
D5 404 ATCCCAAGCTTGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGAGATGATTCTC 463
QY 498 CACCTGTTATCAAGGAATGTTAGAGCATTTTCAACTCAGTTTCCCTACATTAATCTCTG 557
D5 464 CACCCAGTTATCAAGGAATGTTAGATCATTTTCAATTCAGTTCCTCCCAACA---ATCTCTC 520
```


CC orthologue of a thalacress transcription factor isolated from Soybean.
SQ Sequence 969 BP; 311 A; 188 C; 200 G; 270 T; 0 U; 0 Other;

Query Match
Best Local Similarity 49.0%; Score 423; DB 12; Length 969;
Matches 573; Conservative 2; Mismatches 92; Indels 36; Gaps 6;

Qy 81 GGCCTTGACAAATGGAAGAGACTTGTCTGTGATGAATATATTTGCAAAATCATGGGAAG 140
Db 44 GGGCCAGGACAAATGGAAGAGACTTAATCTTGTATCACTATATTTGCCAATCACGGGAAG 103

Qy 141 GTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAACTGAAGGAAAGTTGCCGGC 200
Db 104 GGGTTTGGAACTCTTTGGCCAAAGCTGCTGCACTTAAACGTACCGGAAGAGTTGCCGGC 163

Qy 201 TAAGTGGCTAAATTAATCTCGCTCGTGTAGAGAGGGAATATTACACCGGAGAAC 260
Db 164 TTGGTAGCTAATCTAGCTAGCTGCTGTAGAGAGGGAATATTACACCGGAGAAC 223

Qy 261 AACTTTTGAATATGAGCTCCACGAAAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGC 320
Db 224 AGCTTTTGAATATGAGCTCCACGAAAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGC 283

Qy 321 ATCTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACAAAGGATCCAGAAGC 380
Db 284 ATCTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACAAAGGATCCAGAAGC 343

Qy 381 ACATCAAGCAAGCT--GAGAACTTTTCAGCAACAGAGTAGTAATAATTCTGAGATTAATG 437
Db 344 ACCTCAAGCAAGCTTCCAGCAGCTTCCAGCAACAGAGTAGTAATAATTCTGAGATTAATG 403

Qy 438 ATCAACCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTATTCTC 497
Db 404 ATCCCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTATTCTC 463

Qy 498 CACCTGTTATCAAGGAATGTAGAGCAATTTTCAACTAGTTCCTTACAAATTAATCTTG 557
Db 464 CACCCAGTTATCAAGGAATGTAGATCCATTTTCAATTCAGTTTCCCAACA---AATCCTC 520

Qy 558 ATCAATCCAGTTGTGTACCAATGACAAACAACTAATACTATTGGAGCATGGAGATA 617
Db 521 ATCATTTAGTTGTGTGTACCAATGACAAACAACTAATACTATTGGAGCATGGAGATA 580

Qy 618 GCTGGTCAATG-----CAATTAATGAGGCTGATTAAT-----TA 652
Db 581 TCTGTTCAATGCAATTAATGAGGCTGATTAAT-----TA 640

Qy 653 TTATCAAGATAAAACCTAAGTTTGAAGTTCATTAAGGCTGG--AATGCTTGTGATTA 710
Db 641 ATAATATATATAAACCTAAGTGTCTAAGTTTCCATAAATTAAGCTGTAGTCTCTGGCTTA 700

Qy 711 AACATATTATGGTTTGTATATAGTAGTTGATGTTGG 753
Db 701 AACATGTTA---GGTTGTTTATACAAAGTAGTTGGATGTTGG 740

RESULT 9
ADI43073
ID ADI43073 standard; DNA; 988 BP.
XX
AC ADI43073;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant transcription factor polynucleotide #1008.

XX transgenic; plant; enhanced tolerance to abiotic stress;
KW phosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;

KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
XX transcription factor; gene; ds.
OS Glycine max.
PN US2004019927-A1.
XX
PD 29-JAN-2004.
XX
PP 25-FEB-2003; 2003US-00374780.
XX
PR 18-APR-2001; 2001US-00837944.
XX
PA (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX
PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
XX
PT New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX
PS Claim 1; SEQ ID NO 1536; 435pp; English.
XX
CC The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produce a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC trichomes; reduced apical dominance, reduced trichome density; lack of
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanins, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant in
CC transcription factor, and an orthologue of Arabidopsis thaliana
CC transcription factors isolated in the invention, that can be used in the
CC creation of a transgenic plant with altered traits.
XX
SQ Sequence 988 BP; 325 A; 196 C; 193 G; 274 T; 0 U; 0 Other;

Query Match 46.4%; Score 400.2; DB 12; Length 988;
Best Local Similarity 71.1%; Pred. No. 1e-70;
Matches 594; Conservative 2; Mismatches 210; Indels 29; Gaps 4;

Qy 18 CACCAAGTCATGATATAAACAACACACAGCTGAAGCTCTCAAGATCCTGAAGTGAGAA 77
Db 52 CACACAAAATATATGGACAAAACCA---TGCAACTCATCTCATGATCTCTGAAGTGAGAA 108

the polypeptides encoded by the polynucleotide described above. The transgenic plant is useful for producing a plant that has an altered trait e.g. an enhanced tolerance to abiotic stress (increased tolerance to chilling, germination in cold conditions, freezing tolerance, tolerance to heat, tolerance to drought, tolerance to osmotic stress, tolerance to salt, tolerance to phosphate limitation, tolerance to potassium limitation, decreased sensitivity to nitrogen limitation), altered hormone sensitivity, reduced sensitivity to abscisic acid, an altered response to ethylene, disease resistance, altered susceptibility to Botrytis, altered susceptibility to Fusarium, altered susceptibility to Erysiphe, altered susceptibility to Pseudomonas syringae, altered susceptibility to Sclerotinia, altered sugar sensing, improved seed germination and seedling vigor, early flowering, late flowering, extended period of flowering, an inflorescence architectural change, a change in stem bifurcations, a lack of a shoot meristem, reduced meristem cell differentiation, altered phyllotaxy, altered branching pattern, reduced apical dominance, reduced trichome density, ectopic trichome development, altered trichome development, altered stem morphology, increased root growth, increased root hairs, altered seed development, altered cell proliferation/cell differentiation, premature senescence, delayed senescence, lethality, increased necrosis, an increase in seedling or plant size, decreased plant size, a change in leaf morphology, increased altered leaf development, increased leaf size and mass, glossy leaves, leaf cell expansion, change in seed morphology, altered seed coloration, increased seed size, decreased seed size, altered seed shape, change in leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid content, increased leaf insoluble sugars, decreased leaf insoluble sugars, increased leaf anthocyanins, an alteration of leaf fatty acid content, an alteration of leaf glucosinolate content, change in seed biochemistry, an increase in seed oil content, decrease in seed oil content, increase in seed fatty acid content, decrease in seed fatty acid content, increase in seed protein content, decrease in seed protein content, alteration in seed prenyl lipid content, increase in seed sterols, upregulation of genes involved in secondary metabolism, increase in root anthocyanins, increase in plant anthocyanins, and alteration in light response or shade avoidance. The present sequence encodes an orthologue of a thalecress transcription factor isolated from Soybean.

XX Sequence 988 BP; 325 A; 196 C; 193 G; 274 T; 0 U; 0 Other;

Query Match 46.4%; Score 400.2; DB 12; Length 988;

Best Local Similarity 71.1%; Pred. No. 1e-70;

Matches 594; Conservative 2; Mismatches 210; Indels 29; Gaps 4;

Qy 18 CACAAGCTCAATGGATGATAAAAAACAAGTGTGAAGCGTCTCAAGATCCTCAAGTGAGAA 77
 Db 52 CACAAGCTCAATGGATGATAAAAAACAAGTGTGAAGCGTCTCAAGATCCTCAAGTGAGAA 108
 Qy 78 AAGGCGCTTGACCAATGGAAGAACTTGTATGATGAACATATATTGCAAAATCATGGG 137
 Db 109 AGGACCATGGACCATGGAAGAACTTGTATGATGAACATATATTGCAAAATCATGGG 168
 Qy 138 AAGGTGTTGGAACCTTTTGGCCAAAGCTGCTGCTCAACGTAACGGAAGAGTTGCC 197
 Db 169 AAGGTGTTGGAACCTTTTGGCCAAAGCTGCTGCTCAACGTAACGGAAGAGTTGTC 228
 Qy 198 GGCTAAGGTGCTCAATTAACCTCCGTCCTGATGTTAGAAAGGGAATATTACACCCGAGG 257
 Db 229 GACTCCGTTGGCTAACTACCTTGGTCTGATGTTAGAGAGGAACAATTACACCCGAGG 288
 Qy 258 AACAACTTTTGAATATGAGAGTCCACGCAAAAGTGGGAAACAGGTGGTCCAAAATTGCCA 317
 Db 289 AACAGCTTTTGCATATAGAACTTCAATGCAAAAGTGGGCAATAGGTGGTCCAAAATTGCCA 348
 Qy 318 AGCATCTACCTGGAGGAGCTGATATGATGATCAGAGTATTTGGAGGACAGGATCCAGA 377
 Db 349 AGCATCTTCCAGGAGGAACCTGACATGAGATTAAGAACTTCTGGAGAACAGGATCCAAA 408
 Qy 378 AGCATCTCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTCTGAGATAATG 437
 Db 409 AGCATCTTACAGCAAGCTGAGAACTTTCACACAAACATGTAATTCAGAGAAATATGATCATC 468
 Qy 438 ATCACAAGCTAGCACTAGCCATGTTTCCACCATTGGCTGAGCCCATGGAGATGATTCTTC 497

Db 469 AAGCAAGCACTAGTACTAGCAAAAGTGTCCACCATGGCACATCCAAATGAGACTTTCTCTC 528
 Qy 498 CACCCTGTATCAAGGATGTAGAGCCATTTTCAACTCAGTCCCTACAAATTAATCTCTG 557
 Db 529 CACCCTCATCAAGGAACCTTTTGGCCATTTCCAACTTCCCTACAATCA---CTG 585
 Qy 558 ATCAATCCAGTTGTTTACCAATGACAAACAAACATTAACATTTGGAGCATGGAGGATA 617
 Db 586 ATCAATCAAGTTGTTTACCAACCAACCAACGACAAACAACTATTGGAGCATCGAGGATA 645
 Qy 618 GCTG---GTCATGCAATTAAGTGAACGGTGATTTAA-----ATATT 654
 Db 646 TCTGGTGTCTATGCAATTAAGTGAACGGTGATTTAAACCTAGCTATATGTCATGCTATATA 705
 Qy 655 ATCAAGATAAAACCTAAGTTTGAAGTTCCATAAGGCTGGAATGCTCTVTGGATTAAACA 714
 Db 706 AATCATATATATGATGATATATAAACCTAAGCTCTTGTAGAGTGTGTTCAGGCTTAATAA 765
 Qy 715 TATTATTGGGTTTGTATATATAAGTAGTGTGGATGTTTGGTTCGCTACCAATTAATAGCT 774
 Db 766 CATCATTAGTCTGTTTATATAGTAGTCTAAGTTTGGTGTGTTTGTATGATGATGTGAG 825
 Qy 775 ATGTGCTGTATATATACAGAGATTTTATATATAACTATATCTGCATGCTTTATAT 829
 Db 826 TTAAGAATTAATTTAGTTAGTGGATTGAATATATATAGTAACCTATATATACATCT 880

RESULT 11
 AC007387
 ID AC007387 standard; cDNA; 805 BP.
 XX AC AC007387;
 XX AC AC007387;
 DT 07-AUG-2003 (first entry)
 XX DE
 XX Soybean Myb-related transcription factor #6 cDNA.
 XX Soyben; ss; Myb-related transcription factor; crop improvement;
 XX starch composition; protein composition; plant; gene; grain quality;
 XX disease resistance; pest resistance; oil composition.
 XX Glycine max.
 XX US2003024007-A1.
 XX 30-JAN-2003.
 XX 14-DEC-2001; 2001US-00021811.
 XX 02-DEC-1998; 98US-0110609P.
 XX 01-DEC-1999; 99US-00452244.
 XX (CAHO)/ CAHOON R E.
 XX (ODEL)/ ODELL J T.
 XX Cahoon RE, Odell JT;
 XX PI
 XX MPI: 2003-456317/43.
 XX P-PSDB; ABO1760.
 XX New isolated polynucleotide encoding a Myb-related transcription factor,
 XX useful in plant molecular biology, in particular for the improvement of
 XX crop plants for a variety of traits, including disease and pest
 XX resistance.
 XX Claim 2; Page 37-38; 61pp; English.
 XX The invention relates to an isolated polynucleotide encoding a Myb-
 XX related transcription factor polypeptide. The methods and compositions of
 XX the present invention are useful in plant molecular biology, in
 XX particular for the improvement of crop plants for a variety of traits,
 XX including disease and pest resistance and grain quality improvements such

XX Plant; Myb-related transcription factor; disease resistance; soybean;
 KW Gene; ss.
 XX Glycine max.
 OS US2004040057-A1.
 XX 26-FEB-2004.
 XX 11-SEP-2003; 2003US-00659869.
 XX 02-DEC-1998; 98US-0110609P.
 XX 01-DEC-1999; 99US-00452244.
 XX 14-DEC-2001; 2001US-00021811.
 XX (CAHO/) FAHOON R E.
 XX (FANG/) FANG Y.
 XX (ODEL/) ODELL J T.
 XX (WENG/) WENG Z.
 XX Cahoon RE, Fang Y, Odell JT, Weng Z;
 PI WPI; 2004-238121/22.
 DR P-PSDB; ADJ77735.
 XX Plant Myb transcription factor homologs and their polynucleotide
 PT sequences useful in plant genomic research e.g. identifying nucleotide
 PT fragments that encode a Myb-related transcription factor polypeptide.
 XX Claim 2; SEQ ID NO 33; 62pp; English.
 XX The present invention relates to the isolation of polynucleotide
 CC sequences encoding plant Myb-related transcription factors and the
 CC encoding polypeptide sequences. Also disclosed is a chimeric gene
 CC encoding all or part of a Myb-related transcription factor linked to
 CC suitable regulatory sequences, and a host cell transformed with the
 CC chimeric gene and able to express the chimeric gene resulting in the
 CC production of altered levels of the Myb-related transcription factor in
 CC the transformed cell. The polynucleotide sequences of the invention may
 CC be used to probe a cDNA or genomic library to identify nucleic acid
 CC fragments encoding a Myb-related transcription factor polypeptide and
 CC also for research purposes including generating plants with increased
 CC resistance to disease, etc. The present sequence encodes a plant Myb-
 CC related transcription factor.
 XX Sequence 910 BP; 326 A; 164 C; 170 G; 246 T; 0 U; 4 Other;
 SQ
 Query Match 39.7%; Score 342.6; DB 12; Length 910;
 Best Local Similarity 75.8%; Pred. No. 3.4e-59;
 Matches 498; Conservative 1; Mismatches 130; Indels 28; Gaps 5;
 QY 8 CTCTATCACACACAAGTCAATGATGATAAAACAAACAGTGAAGCTCTCAAGATCCT 67
 DB 16 CTCTAGCGTGACACAAAATAATGACAAAAACCA--TGGGACTCATCTCATGATCCA 72
 QY 68 GAAGTGAGAAAGCGCTTGGACATGGAAGACTTGCATCTTGAATATATTGCA 127
 DB 73 GAAGTGAGAAAGGGACCATGGGATGATGGAAGAAGACTTGCATCTTGAATATATTGCA 132
 QY 128 AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAACGTAACGGA 187
 DB 133 AATCAGGTGAAGGTGTTTGGAACTCTTTAGCCAAAGCTTCTGGTCTTAAACGACGGA 192
 QY 188 AAGAGTTGCCGCTAAGTGGGTAAATTAACCTTCCTGATGTTAGAGAGGAATATT 247
 DB 193 AAGAGTTGTCGACTCGTGGCTAAACTACCTTCCTGATGTTAGAGAGGAACATTT 252
 QY 248 ACACCGGAGAACACTTTGATTTGGAGCTCCAGCAAGTGGGGAACAGGTGTC 307
 DB 253 ACACCGGAGAACAGCTTTGATTTGATAGAACTTCATGCAAAAGTGGGGCAATAGGTGGTCC 312
 QY 308 AAAATTGCCAAGCATCTACCTGGAGGACTGATATATGATCAAGAACTATTGGAGGACA 367

DB 313 AAAATTGCAAGCATCTTCCAGGAAGAACTGACAATGAGATTAGAACTTCTTGGAAGACT 372
 QY 368 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGT----- 418
 DB 373 AGGATCCAGAAGCACATTAAGCAAGCTGAGAACTTCAACAACAACTGGTAATTTCATCAGAG 432
 QY 419 AATAAATTCGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAG 478
 DB 433 AATAGTAATTAATGATCATCAAGCAAGCAATAGCAAGGTGTCCACCATGGCACAT 492
 QY 479 CCCATGGAGATGTAATCTCCACCTGTTATCAAGGAATGTAGAGCCATTTTCAACTCAG 538
 DB 493 CCAAAATGAGACTTCTCTTCCACCTCATACCAAGCAACTTTTGAGCCATTTCAACCTCAA 552
 QY 539 TTCCCTACAATTAATCCTGATCAATCCAGTGTGTGTACCAATGACAAACAACATTAAAC 598
 DB 553 TTC-----CTACAATCAATGATCAATCAAGTTGTTGTAC-----CAGCAACAACAAC 599
 QY 599 TATTGGAGCATGGAGGATAGCTG---GTCAATGCAATTACTGAAACGGTGAATTAATA 652
 DB 600 TATTGGAGCATGGAGGATATCTGGTCTGCTATGCAATTACTCAATGGAGATWAATTA 656
 RESULT 15
 ACID07384
 ID ACD07384 standard; cDNA; 530 BP.
 XX AC ACD07384;
 XX 07-AUG-2003 (first entry)
 XX Soybean Myb-related transcription factor #3 cDNA.
 XX Soyben; ss; Myb-related transcription factor; crop improvement;
 KW starch composition; protein composition; plant; gene; grain quality;
 KW disease resistance; pest resistance; oil composition.
 XX Glycine max.
 XX US2003024007-A1.
 XX 30-JAN-2003.
 XX 14-DEC-2001; 2001US-00021811.
 XX 02-DEC-1998; 98US-0110609P.
 XX 01-DEC-1999; 99US-00452244.
 XX (CAHO/) FAHOON R E.
 XX (ODEL/) ODELL J T.
 XX Cahoon RE, Odell JT;
 PI WPI; 2003-456317/43.
 DR P-PSDB; ABO01757.
 XX New isolated polynucleotide encoding a Myb-related transcription factor,
 PT useful in plant molecular biology, in particular for the improvement of
 PT crop plants for a variety of traits, including disease and pest
 PT resistance.
 XX Claim 2; Page 34; 61pp; English.
 XX The invention relates to an isolated polynucleotide encoding a Myb-
 CC related transcription factor polypeptide. The methods and compositions of
 CC the present invention are useful in plant molecular biology, in
 CC particular for the improvement of crop plants for a variety of traits,
 CC including disease and pest resistance and grain quality improvements such
 CC as oil, starch or protein composition. The present sequence represents
 CC cDNA encoding a soybean Myb-related transcription factor
 XX Sequence 530 BP; 183 A; 111 C; 110 G; 116 T; 0 U; 10 Other;

Query Match		35.2%;	Score 303.8;	DB 8;	Length 530;	
Best Local Similarity		82.6%;	Pred. No. 1.9e-51;			
Matches 369;		Conservative	0;	Mismatches 76;	Indels 2; Gaps 2;	
Qy	36	AAAAACAACAGCTGTAAGACGCTCTCAAGATCCTGGAAGTGAGAAAAGGGCCCTTGGACAATGG				95
Db	11					70
Qy	96	AAGAAGACTTCGATCTTGATGAACCTATATTGCAAAATCATGGGGAAGGTGTTTGGAACTCTT				155
Db	71					130
Qy	156	TGGCCAAAGCTGCTGGTCTCAAACTAAACGGAAAGAGTTGCCGGCTAAGGTGGCTAAATT				215
Db	131					190
Qy	216	ACCTCCGTCCTGATGTTAGAAAGGGAATATTACACCGAGGAACAACCTTTTGATTATGG				275
Db	191					250
Qy	276	AGCTCCACGCAAGTGGGGAAACAGGTGTCCTCAAAATTTGCCAAGCATCTACCTGGAAGGA				335
Db	251					310
Qy	336	CTGATAATGAGATCAAGAACTATTTTGGAGGACAAGGATCCAGAAGCACATCAAGCAAGCTG				395
Db	311					369
Qy	396	AGAACTTTTCAGCAACAGAGTAGTAATAATTTCTGAGATAAATGATCACCAAGCTAGCACTA				455
Db	370					428
Qy	456	GCCATGTTTCCACCATGGCTGAGCCCA				482
Db	429					455

Search completed: September 28, 2005, 09:29:17
Job time : 591 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 08:42:38 ; Search time 193 Seconds
(without alignments)
7316.616 Million cell updates/sec

Title: US-10-659-869A-35
Perfect score: 863
Sequence: 1 gcacgagctctatcacacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248.4	28.8	862	4	US-09-640-211A-2076
2	223	25.8	907	4	US-09-640-211A-2077
3	215	24.9	1033	4	US-09-533-029-87
4	169.4	19.6	1347	4	US-09-533-029-39
5	165.2	19.1	673	4	US-09-640-211A-328
6	161.6	18.7	389	4	US-09-640-211A-283
7	161.6	18.7	417	4	US-09-640-211A-1496
8	152.4	17.7	524	4	US-09-640-211A-1934
9	132.8	15.4	516	4	US-09-640-211A-483
10	132.8	15.4	516	4	US-09-640-211A-588
11	131.4	15.2	542	4	US-09-640-211A-565
12	130.8	15.2	918	4	US-09-533-029-93
13	130.6	15.1	389	4	US-09-640-211A-1559
14	123.2	14.3	373	4	US-09-640-211A-1967
15	123	14.3	378	4	US-09-640-211A-282
16	121.6	14.1	1157	4	US-09-453-387A-3
17	120.6	14.0	1678	4	US-09-640-211A-2074
18	120.4	14.0	1006	4	US-09-453-387A-1
19	118.4	13.7	404	4	US-09-640-211A-1497
20	118	13.7	526	4	US-09-640-211A-1985
21	117.2	13.6	476	4	US-09-640-211A-1987
22	117.2	13.6	631	4	US-09-640-211A-1718
23	117.2	13.6	1150	4	US-09-640-211A-2088
24	116.8	13.5	360	4	US-09-640-211A-1781
25	116.2	13.5	1081	4	US-09-453-387A-5
26	116	13.4	420	4	US-09-640-211A-494
27	116	13.4	424	4	US-09-640-211A-1965

ALIGNMENTS

RESULT 1
US-09-640-211A-2076
; Sequence 2076, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2076
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-640-211A-2076

Query Match	28.8%	Score 248.4;	DB 4;	Length 862;
Best Local Similarity	72.1%	Pred. No. 7.7e-52;		
Matches	356;	Conservative	0;	Mismatches 126;
				Indels 12;
				Gaps 2;
Qy	8	CTCTATCACACACACAAGTCAATGGATATAAAGAGAGTGTAAAGAGCTCTCAAGATCCT	67	
Db	16	CTCTCCAAAGCTGAACATGGACAGAACCCAGACGACAGTGGTAAAGTCCCAAGATGTC	75	
Qy	68	GAAGTGAAGAAAGGGCTTGGACAAATGGAGAGAGACTTGATCTTGATGAACATATTTGCA	127	
Db	76	GAGTGAAGAAAGGGCGCTGGACGATGGAAGAGATCTCATCTCATCACTCAATATAGCG	135	
Qy	128	AATCATCGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGTGTCTCAAAACGTAACGGA	187	
Db	136	ATTCACGGCGAAGGCGAGTGTGGAACTCCCTAGCCAAAGCTGTGTCTTAAACGTTACCGGG	195	
Qy	188	AAGAGTTGCCGGCTAAGTGGCTAAATTAACCTCCGTCCTGATGTTAGAAGAGGAATATT	247	
Db	196	AAGAGTTGTCCGCTCCGCTGGCTGAATCTATCTGCGACCCGACGCTCGGAGAGGCAACATC	255	
Qy	248	ACACCCGAGGACAACTTTTGGATATGAGCTCCACCAAGTGGGGAACAGGTGTCTCC	307	
Db	256	ACTACTGAGGAGCAGCTCTCTGATCATGGAATGATGATGAGTGGGGAACAGGTGTCT	315	
Qy	308	AAATTCGCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGTGAGGACA	367	
Db	316	AAATTCGCAAGCATCTTCCGGAAGGACTGATATGAGATCAAGAACTATTGTGAGGACT	375	


```
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G239
US-09-533-029-39

Query Match      19.6%; Score 169.4; DB 4; Length 1347;
Best Local Similarity 69.5%; Pred. No. 2.9e-32;
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 60 AGATCTCTGAAGTGAAGAGGCGCTTGGACAATCGAAGAGACTTGATCTTGATGAAC 119
DB 47 AAGATCTCTGATGACGGAAGGTCCATGACCGGAGGAAGATGCAATCCTAGTCAACT 106

QY 120 ATATTGCCAATCATGGGAAGTGTGGAACCTTTTGGCCAAAGCTGCTGCTCTCAAC 179
DB 107 TCGTCTCTATTATCGCGATGCTCGTTGGAACCAACATCGCTCGTTCTCTGGGCTAAAGC 166

QY 180 GTAACGGAAGAGTTCGCGCTTAAGTGGCTTAATTAATCTCGTCTCTGATGTTAGAAGAG 239
DB 167 GAATCTGGAAGTGTAGATTAAGATGCTTAATTAATCTAGTCCAGATGTTAGAAGAG 226

QY 240 GGAATATTACACCCGAGGAACAACATTTTGAATATGAGCTCCAGCAAAAGTGGGGAACA 299
DB 227 GCAACATCACTCTCGAAGAACAAATTAATGATCTCTCAAACTCATTTCTTTGGGSCAATA 286

QY 300 GGTGCTCAAAATTCGCAAGCATCTACCTGGAGGACTGATAATGAGATCAAGAACTATT 359
DB 287 GGTGCTGAAGATTGCGCAATATCTACCGGGAAGAACAGATAATGAATAAAGAAATTATT 346

QY 360 GGAGGACAAGGATCCAGAGACACATCAAGCA 390
DB 347 GGAGAACTCGAGTCCAAAAGCAAGCAAAACA 377

RESULT 5
US-09-640-211A-328
; Sequence 328, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 328
; LENGTH: 673
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-328

Query Match      19.1%; Score 165.2; DB 4; Length 673;
Best Local Similarity 70.4%; Pred. No. 2.5e-31;
Matches 221; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 77 AAAGGCGCTTGACAATCGAAGAGACTTGATCTTGATGAACACTATATTGGCAATCATGGG 136
DB 2 AGAGTCCGTGACGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 61

; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G239
US-09-533-029-39

Query Match      19.6%; Score 169.4; DB 4; Length 1347;
Best Local Similarity 69.5%; Pred. No. 2.9e-32;
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 60 AGATCTCTGAAGTGAAGAGGCGCTTGGACAATCGAAGAGACTTGATCTTGATGAAC 119
DB 47 AAGATCTCTGATGACGGAAGGTCCATGACCGGAGGAAGATGCAATCCTAGTCAACT 106

QY 120 ATATTGCCAATCATGGGAAGTGTGGAACCTTTTGGCCAAAGCTGCTGCTCTCAAC 179
DB 107 TCGTCTCTATTATCGCGATGCTCGTTGGAACCAACATCGCTCGTTCTCTGGGCTAAAGC 166

QY 180 GTAACGGAAGAGTTCGCGCTTAAGTGGCTTAATTAATCTCGTCTCTGATGTTAGAAGAG 239
DB 167 GAATCTGGAAGTGTAGATTAAGATGCTTAATTAATCTAGTCCAGATGTTAGAAGAG 226

QY 240 GGAATATTACACCCGAGGAACAACATTTTGAATATGAGCTCCAGCAAAAGTGGGGAACA 299
DB 227 GCAACATCACTCTCGAAGAACAAATTAATGATCTCTCAAACTCATTTCTTTGGGSCAATA 286

QY 300 GGTGCTCAAAATTCGCAAGCATCTACCTGGAGGACTGATAATGAGATCAAGAACTATT 359
DB 287 GGTGCTGAAGATTGCGCAATATCTACCGGGAAGAACAGATAATGAATAAAGAAATTATT 346

QY 360 GGAGGACAAGGATCCAGAGACACATCAAGCA 390
DB 347 GGAGAACTCGAGTCCAAAAGCAAGCAAAACA 377

RESULT 6
US-09-640-211A-283
; Sequence 283, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-283

Query Match      18.7%; Score 161.6; DB 4; Length 389;
Best Local Similarity 71.6%; Pred. No. 1.7e-30;
Matches 212; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 8 CTCTATCACACACACAAGTCAATGGATATAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 67
DB 29 CTCTCAAGCTGAACATGGACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 88

QY 68 GAAGTGAGAAAGGGCGCTTGGACAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127
DB 89 GAGGTGAGAAAGGGCGCTTGGACAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 148

QY 128 AATCATGCGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAAAAGTAAACGGA 187
DB 149 AATCAGCGGAGAGGAGTGGAACTCTCTAGCCAAAGCTGCTGCTCTTAAACGAGTACCGGG 208

QY 188 AAGAGTTGCCGGCTTAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 247
DB 209 AAGAGTTGCTGGCTCCGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268

QY 248 ACACCCGAGGAGCAACTTTTGAATATGAGGCTCACCGCAAGAGTGGGGAAGAGAGAGAG 303
DB 269 ACTACTGAGGAGGAGCTCTCTGATCATGGAAGTGCATGCAAGTGGGGAAGAGAGAGAG 324

RESULT 7
US-09-640-211A-1496
; Sequence 1496, Application US/09640211A
; Patent No. 6833446
```

```
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1496
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1496

Query Match      18.7%; Score 161.6; DB 4; Length 417;
Best Local Similarity 71.6%; Pred. No. 1.7e-30;
Matches 212; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY      8  CTCATCACACACACACAGTCAATGGATATAAATAACACAGTGTAAAGCGTCTCAAGATCCT 67
Db      29  CTCCTCAAGCTGAACATGACAGAGAGAGCCAGACGACGAGTGGTAAGTCCCAAGATGTC 88

QY      68  GAAAGTGAGAAAGGGCGCTTGGACAATGGACAATGGAGAGACTTGATCTTGATGAACATATATTGCA 127
Db      89  GAGGTGAGAAAGGGCGCTGGACGATGGAGAGGATCTCATCTCATCACTACATAGGG 148

QY     128  AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGGTCTCAAAAGCTAACCGGA 187
Db     149  AATCAGCGGGAAGGAGTGGAACTCCCTAGCCAAAGCTGCTGGTCTTAAAGCTACCGGG 208

QY     188  AAGAGTTCGGCTAAGGTGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 247
Db     209  AAGAGTTCGGCTCGGTGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 268

QY     248  ACACCGGAGGACAACTTTTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 303
Db     269  ACTACTGAGGAGCAGCTCTTGATCATGGAATGCATGCAAGTGGGGAACAGGTG 324

RESULT 8
US-09-640-211A-1934
; Sequence 1934, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1934
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1934

Query Match      17.7%; Score 152.4; DB 4; Length 524;
Best Local Similarity 67.1%; Pred. No. 3.4e-28;
Matches 216; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY      60  AAGATCTCTGAAGTGAGAAAAGGGCTTGGACAAATGGAAAGAGACTTGATCTTGATGAAGT 119
Db     179  AAGAGGTGACCTCGAAAGGGGCGCATGGATCTCGGAGGAAGACAAATTTGCTCATTTCACT 238

QY     120  ATATTGCAAAATCATGGGAAGGTGTTTGGAACTCTTTGGCCAAAAGCTGCTGGTCTCAAAAC 179
Db     239  CGATCACATGTCACGGCGAGGGACGCTGGAATATGTTGGCGAAGAGCGCAGGATTTGAAGA 298

QY     180  GTAACGGNAAGAGTTCGGGCTAAGGTGGCTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 239
Db     299  GAACTGGCAAAAGCTGCAGATTAAGGTGGCTGAATTAATTAATTAATTAATTAATTAATTAATTA 358

QY     240  GGAATATTATACACCGGAGGAACAACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 299
Db     359  GGAATCTCACCCCGCAAGAACAGCTCATGATCTTGAATTTCAACACAAATGGGCAACA 418

QY     300  GGTGTTCCAAATTTGCCAAGCATCTACCTGGAAGGACTGATTAATGAGATCAAGAACTATT 359
Db     419  GGTGTTCCAAATTCGCGCAGTATCTCCAGGAAGGACAGATAACGAGATCAAGAACTACT 478

QY     360  GGAGGACAAGGATCCAGAGCA 381
Db     479  GGAGGACGCGGTGCAGAAAGCA 500

RESULT 9
US-09-640-211A-483
; Sequence 483, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-483

Query Match      15.4%; Score 132.8; DB 4; Length 516;
Best Local Similarity 67.1%; Pred. No. 2.4e-23;
Matches 188; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY      71  GTGAGAAAAGGGCTTGGACAATGGAAAGAGACTTGATCTTGATGAACATATATTGCAAT 130
Db     237  GTGAGGAGGGGACCTTGGACTGTTGATGAGGACATGAGCCTTATTTCGATGCGTAACCCACC 296

QY     131  CATGGGAAGGTGTTTGGAACTCTTTGGCCAAAAGCTGCTGGTCTCAAAAGTAAACGGAAG 190
Db     297  CGGGTGAAGTTCGATGGAACACAGTAGCCAAATTTGCAGGGCTAAAGAGACAGAGAAAG 356

QY     191  AGTTGCCGGCTAAGTGGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 250
Db     357  AGTGCAGATTGAGATGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 416

QY     251  CCCGAGGAACAACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 310
Db     417  CCGGAAGAGCAGCTATTAAATCTTTGAATCCACCGTCTCTGGGGTAAACAGATGGTCCAAG 476

QY     311  ATTGCCAAGCATCTACCTGGAAGGACTGATTAATGAGATCA 350
Db     477  ATTGCACGGCAACTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516

RESULT 10
```

```
US-09-640-211A-588
; Sequence 588, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 588
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-640-211A-588

Query Match 15.4%; Score 132.8; DB 4; Length 516;
Best Local Similarity 67.1%; Pred. No. 2.4e-23;
Matches 188; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 71 GTGAGAAAGGCGCTTGGACAATGGAAGAGCTTGATCTTGTGTAAGAACTATATTGCAAAAT 130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
237 GTGAGGAGGAGCCTTGGACCTGTTGATGAGGACATGAGCCCTATTTCGATGCGTAACCAAC 296
Qy 131 CATGGGGAAGGTGTTTGAACCTCTTGGCCAAAGCTGTCTCAAAACGTAACGGAAAG 190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 CGGGGTGAAGTTCGATGGAACACAGTAGTACCAAAATTTGCAGGGCTAAAGAGAACAGGAAAG 356
Qy 191 AGTTCCCGCTAAAGTGTGCTAAATTTACCTCGTCTGATGTTAGAGAGGGAATATTACA 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
357 AGTGCAGATTGAGATGCTTAATATCTTCGCCCGGATGTTAAACGTGGAAACATTAACG 416
Qy 251 CCCGAGGAACAACTTTTGTATTATGAGCTCCACGCAAAAGTGGGAAACAGGTGTCACAA 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 CCGGAAGAGCAGCTATTAACTCTTGAACCTCCACCGTCTCTGGGGTAAACAGATGTTCCAAG 476
Qy 311 ATTGCCAAGCATCTACCTGGAGAGCTGATTAATGAGATCA 350
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 ATTGCAGGCAACTCCCGAGGAGGAGTGAACACGAAATCA 516

RESULT 11
; Sequence 565, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-640-211A-565

Query Match 15.2%; Score 131.4; DB 4; Length 542;
Best Local Similarity 65.6%; Pred. No. 5.3e-23;
Matches 208; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

Qy 60 AAGATCCTGAAATGAGAAAGGCGCTTGGACAATGGAAGAGACTTGATCTTGTGAACT 119

US-09-640-211A-588
; Sequence 588, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 588
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-640-211A-588

Query Match 15.4%; Score 132.8; DB 4; Length 516;
Best Local Similarity 67.1%; Pred. No. 2.4e-23;
Matches 188; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 71 GTGAGAAAGGCGCTTGGACAATGGAAGAGCTTGATCTTGTGTAAGAACTATATTGCAAAAT 130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 ATATTGCAAAATCATGGGGAAGGTGTTTGGAACTCTTTTGGCCAAAGCTGTGCTCTCAAAAC 179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
136 TTGTAATTTTGCACGGAGAGGAGCGTGGAACTTTTCGCCAGAGCATCTGGCCTCCAGA 195
Qy 180 GTAACGGAAGAGTTCGCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGTATGTAGAAAG 239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 GAACTGGTAAGAGCTGCCGCTAAGGTGGGTAACTATCTCCGCGCTGATCTCAAGCGGA 255
Qy 240 GGAATATTACACCGGAGGAACAATTTTGAATATTGAGCTCCAGCAAGATGGGGAACA 299
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 GCAAGATCACTCTCTGAAGAAGACGTTTGAATATTGAATTCATCGCCGTTGGGGAATA 315
Qy 300 GGTGGTCCAAAATTTGCCAAGCATCTACCTGGAAGGAGTGAATAGATCAAGAACTATT 359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 GGTGGTCTCGTATTGCACAAAGTTTACCGGAGGAGGACAAATGAATCAAGAAATTTCT 375
Qy 360 GGAGGACAAAGATCCAG 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 GGAGAACTCGTATGAAG 392

RESULT 12
US-09-533-029-93
; Sequence 93, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G255
; US-09-533-029-93

Query Match 15.2%; Score 130.8; DB 4; Length 918;
Best Local Similarity 63.9%; Pred. No. 8.9e-23;
Matches 198; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 72 TGAGAAAGGCGCTTGGACAATGGAAGAGACTTGATCTTGTGAACTATATTGCAAAATC 131
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 TGACAAAGGTGCTTGGACTTAAGAGAGAGATGAGAGACTAGTCTCTTACATCAAGTCTC 123
Qy 132 ATGGGGAAGGTGTTTGAACCTCTTTGGCCAAAGCTGTGCTCAAAACGTAACGGAAGA 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 ACGGTGAAGGTGTTGTCGATCTCTTCTAGAGCGCGTGGTCTCTCTTCGCTGCGGTAAA 183
Qy 192 GTTCCCGGCTAAGGTGCTAAATTAACCTCGTCTGATGTTAGAGAGGGAATATTACAC 251
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 GCTGCCGCTCTTCGGTGGATTAACTATCTCCGACCTGATCTCAAAAGAGGAACTTTTACAC 243
```

Qy	252	CCGAGGAACAACTTTTGATTATGGAGTCCACGAAAGTGGGGAAACAGTGGTCCAAA	311
Db	244	ATGATGAAGATGAACCTTATCATCAAGCTTCATAGCCTCTAGGCAACAAGTGGTCTTTGA	303
Qy	312	TTGCCAAGCATCTACTCTGGAGGACTGATAATAGATCAAGAAGTATTGGAGGACAAGGA	371
Db	304	TTGGCGCGAGATTACTCTGGAGAAACAGATAACGAGATCAAGAAGTACTGTGGAACACACATA	363
Qy	372	TCCAGAAAGCA	381
Db	364	TAAAGAGGAA	373

```

RESULT 13
US-09-640-211A-1559
; Sequence 1559, Application US/09640211A
; Patent NO. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1559
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1559

```

```

; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; ; TITLE OF INVENTION: Modification of Gene Transcription
; ; FILE REFERENCE: 11000.1021C1U
; ; CURRENT APPLICATION NUMBER: US/09/640,211A
; ; CURRENT FILING DATE: 2000-08-16
; ; NUMBER OF SEQ ID NOS: 2368
; ; SOFTWARE: FastSeq for Windows Version 4.0
; ; SEQ ID NO 1967
; ; LENGTH: 373
; ; TYPE: DNA
; ; ORGANISM: Eucalyptus grandis
US-09-640-211A-1967

```

```
QY 144 TTGTGAACCTCTTTGGCCAAAGCTGTGTCTCAAAACGTAAACGGAAGAGTTGCCGGCTAA 203
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 61 GTTGGCGGGCTGTTCCCAAGCTTGTGACTGTGCGGTGTGGAAGAGTTGCAGGCTGA 120
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 204 GGTGGCTAAATTACCTCGTCTCTGATGTTAGAGAGGGAATATTACACCCCGAGGAACAAC 263
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 121 GGTGACCAATTACTGAGGCCAGACTTGAAGAGAGGCCCTTTGTCCGAGTATGAAGAGA 180
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 264 TTTTGATTATGGAGCTCCACGCAAAAGTGGGGAACAGGTGTCCAAAATTGCCAAGCATC 323
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 181 AAATGGTCATTGACCTCCATCGCAACTTGGCAACAGATGGTCGAAAATAGCCTCTCACC 240
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 324 TACCTGGAGGACTGATATGAGATCAAGAATCTTTGGAGGACAAGGATCCAGAAGCACA 383
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 241 TCCCGGGAAGAACAGACAATGAGATCAAGAATCACTGGGAACACTCACATCAAGAAGAAGC 300
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 384 TCAAGCA 390
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 301 TCAAGAA 307
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
```

Search completed: September 28, 2005, 11:40:41
Job time : 194 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 10:38:02 ; Search time 741 Seconds
(without alignments)
7787.119 Million cell updates/sec

Title: US-10-659-869A-35
Perfect score: 863
Sequence: 1 gcacgagcttattcacacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7400732 seqs, 3343137571 residues

Total number of hits satisfying chosen parameters: 14801464

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	861.8	99.9	863	14	US-10-021-811-35
2	861.8	99.9	863	18	US-10-659-869-35
3	824.4	95.5	947	17	US-10-425-114-10661
4	800.4	92.7	968	18	US-10-374-780A-1537
5	800.4	92.7	968	18	US-10-412-699B-1631
6	610.2	70.7	988	14	US-10-021-811-29
7	610.2	70.7	988	18	US-10-659-869-29

8	471	54.6	1097	18	US-10-424-599-85949	Sequence 85949, A
9	423	49.0	969	17	US-10-374-780A-1538	Sequence 1538, Ap
10	423	49.0	969	18	US-10-412-699B-1632	Sequence 1632, Ap
11	409.6	47.5	1094	18	US-10-424-599-126212	Sequence 126212, A
12	406.2	47.1	879	18	US-10-425-114-10688	Sequence 10688, A
13	400.2	46.4	988	17	US-10-374-780A-1536	Sequence 1536, Ap
14	400.2	46.4	988	18	US-10-412-699B-1630	Sequence 1630, Ap
15	399.8	46.3	877	18	US-10-425-114-10895	Sequence 10895, A
16	394.4	45.7	805	14	US-10-021-811-37	Sequence 37, Appl
17	394.4	45.7	805	18	US-10-659-869-37	Sequence 37, Appl
18	342.6	39.7	910	14	US-10-021-811-33	Sequence 33, Appl
19	342.6	39.7	910	18	US-10-659-869-33	Sequence 33, Appl
20	303.8	35.2	530	14	US-10-021-811-31	Sequence 31, Appl
21	303.8	35.2	530	18	US-10-659-869-31	Sequence 31, Appl
22	298.8	34.6	392	11	US-09-922-293-3810	Sequence 3810, Ap
23	287.4	33.3	557	14	US-10-021-811-27	Sequence 27, Appl
24	287.4	33.3	557	18	US-10-659-869-27	Sequence 27, Appl
25	273.6	31.7	681	21	US-10-487-901-3444	Sequence 3444, Ap
26	265.4	30.8	681	17	US-10-341-961A-197	Sequence 197, App
27	250.8	29.1	462	21	US-10-487-901-6907	Sequence 6907, App
28	248.4	28.8	862	20	US-10-856-499-2076	Sequence 2076, Ap
29	234.6	27.2	731	17	US-10-374-780A-1981	Sequence 1981, Ap
30	234.6	27.2	731	18	US-10-412-699B-173	Sequence 173, App
31	234.6	27.2	731	18	US-10-412-699B-1757	Sequence 1757, Ap
32	223	25.8	907	20	US-10-856-499-2077	Sequence 2077, Ap
33	221.4	25.7	506	18	US-10-424-599-68025	Sequence 68025, A
34	215	24.9	1033	10	US-09-533-029-87	Sequence 87, Appl
35	215	24.9	1033	15	US-10-278-536-85	Sequence 85, Appl
36	215	24.9	1033	17	US-10-225-066A-451	Sequence 451, App
37	215	24.9	1033	17	US-10-374-780A-1969	Sequence 1969, Ap
38	215	24.9	1033	17	US-10-412-699B-1745	Sequence 1745, Ap
39	215	24.9	1033	22	US-10-225-066A-451	Sequence 451, App
40	209	24.2	235	11	US-09-922-293-2831	Sequence 2831, Ap
41	207.8	24.1	291	18	US-10-424-599-26591	Sequence 26591, A
42	206.4	23.9	1011	17	US-10-225-068-1	Sequence 1, Appli
43	206.4	23.9	1011	17	US-10-225-068-21	Sequence 21, Appl
44	206.4	23.9	1011	17	US-10-225-068A-229	Sequence 229, App
45	206.4	23.9	1011	17	US-10-302-267-129	Sequence 129, App

ALIGNMENTS

RESULT 1
US-10-021-811-35
; Sequence 35, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION: Rebecca E.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Glycine max
US-10-021-811-35

Query Match 99.9%; Score 861.8; DB 14; Length 863;
Best local similarity 100.0%; Pred. No. 2.2e-194;
Matches 863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGCTTATCACACACACAGTCAATGGATATAAAACACAGTGAAGCTCTCA 60
Db 1 GCACGAGCTTATCACACACACAGTCAATGGATATAAAACACAGTGAAGCTCTCA 60

RESULT 3
US-10-425-114-10661
; Sequence 10661, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10661
; LENGTH: 947
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700957658_FLI
US-10-425-114-10661

Query Match 95.5%; Score 824.4; DB 18; Length 947;
Best Local Similarity 99.2%; Pred. No. 1.8e-185;
Matches 825; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 CTCATCACACACACAAAGTCAATGGATGATAAAAAACACAGGTGAAGAGCTCTCAAGATCCT 67
DB 27 CTCATCACACACACAAAGTCAATGGATGATAAAAAACACAGGTGAAGAGCTCTCAAGATCCT 86

QY 68 GAAGTGAAGAAAGGCGCTTGACAAATGGAAAGAGACTTGATCTTGATGAACATATATGCA 127
DB 87 GAAGTGAAGAAAGGCGCTTGACAAATGGAAAGAGACTTGATCTTGATGAACATATATGCA 146

QY 128 AATCATGGGAAGAGTGTTCGAACTCTTTGCGCAAGCTCTGCTCTCAACAGTAAACGGA 187
DB 147 AATCATGGGAAGAGTGTTCGAACTCTTTGCGCAAGCTCTGCTCTCAACAGTAAACGGA 206

QY 188 AAGATGTCGCGCTAAGGTGGCTAAATACCTCGTCTCGTCTGATGTTAGAAGAGGGAATATT 247
DB 207 AAGATGTCGCGCTAAGGTGGCTAAATACCTCGTCTCGTCTGATGTTAGAAGAGGGAATATT 266

QY 248 ACACCGAGGAACAACCTTTTGTATATGGAGCTCCACCAAGTGGGGGAAAACAGGTGGTCC 307
DB 267 ACACCGAGGAACAACCTTTTGTATATGGAGCTCCACCAAGTGGGGGAAAACAGGTGGTCC 326

QY 308 AAAATTGCCAAGCATCTACCTGGAAGACTGATAATGAGATCAAGAACTATTGGAGGACA 367
DB 327 AAAATTGCCAAGCATCTACCTGGAAGACTGATAATGAGATCAAGAACTATTGGAGGACA 386

QY 368 AGGATCCAGAAGCACATCAAGCAAGCTGAGAATCTTTTCAGCAACAGAGTAGTAATAATTCT 427
DB 387 AGGATCCAGAAGCACATCAAGCAAGCTGAGAATCTTTTCAGCAACAGAGTAGTAATAATTCT 446

QY 428 GAGATAAATGATCACCAGCTAGCATGAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 487
DB 447 GAGATAAATGATCACCAGCTAGCATGAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 506

QY 488 ATGTATTCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTTACA 547
DB 507 ATGTATTCTCCACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTTACA 566

QY 548 ATTAATCCTGATCAATCCAGTTGTTGTACCAATGACAAACACATTAATTAATTGGAGC 607
DB 567 ATTAATCCTGATCAATCCAGTTGTTGTACCAATGACAAACACATTAATTAATTGGAGC 626

QY 608 ATGAGGATAGCTGGTCAATGCAATTAATCTGAAACGGTGATTAATATATCAAGATAAAC 667

DB 627 ATCGAGGATAGCTGGTCAATGCAATTAATCAACGGTGATTAATAATATATCAAGATAAAC 686
QY 668 CTAAGTTTGAAGTTCATTAAGCTGGAATGTCTTGTGATTAAACATATATTATGGGTTT 727
DB 687 CTAAGTTTGAAGTTCATTAAGCTGGAATGTCTTGTGATTAAACATATATTATGGGTTT 746
QY 728 GTTTATATAAGTAGTAGTGGATGTTTGGTTTTCGCTACCATTTAGCTATGCTGTGAATA 787
DB 747 GTTTATATAAGTAGTAGTGGATGTTTGGTTTTCGCTACCATTTAGCTATGCTGTGAATA 806
QY 788 TATACGAGATTTATATATAAATCAATATCTGATGCTTTATATATAAAAAAAA 839
DB 807 TATACGAGATCTTATATATAAATCAATATCTGATGCTTTATATATAATTTTAAA 858

RESULT 4
US-10-374-780A-1537
; Sequence 1537, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James E
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1537
; LENGTH: 968
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1322
US-10-374-780A-1537

Query Match 92.7%; Score 800.4; DB 17; Length 968;
Best Local Similarity 97.4%; Pred. No. 9.1e-180;
Matches 810; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

QY 8 CTCATCACACACACAAAGTCAATGGATGATAAAAAACACAGGTGAAGAGCTCTCAAGATCCT 67

Db 51 CTCTATCACACACACAGTCAATGGATAAAAAACAACGTGTGCAACACGCTCTCAAGATCCT 110
 Qy 68 GAAGTGAGAAAGGGCCTTGACAAATGGAAGAGCTTGATCTTGATCAACTATATGCA 127
 Db 111 GAAGTGAGAAAGGGCCTTGACAAATGGAAGAGCTTGATCTTGATCAACTATATGCA 170
 Qy 128 AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACAGCTAAAGGA 187
 Db 171 AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACAGCTAAAGGA 230
 Qy 188 AAGAGTTCGGGCTAAGGTGGCTAAATTAATCTCGCTCTGATGTAGAGAGGAATATT 247
 Db 231 AAGAGTTCGGGCTAAGGTGGCTAAATTAATCTCGCTCTGATGTAGAGAGGAATATT 290
 Qy 248 ACACCCGAGGAGCAACTTTTGAATATGAGAGCTCAACGAAAGTGGGAAACAGGTGTCC 307
 Db 291 ACACCCGAGGAGCAACTTTTGAATATGAGAGCTCAACGAAAGTGGGAAACAGGTGTCC 350
 Qy 308 AAAATTGCGAAGCACTTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACA 367
 Db 351 AAAATTGCGAAGCACTTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACA 410
 Qy 368 AGGATCCAGAAGCAGATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATTTCT 427
 Db 411 AGGATCCAGAAGCAGATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATTTCT 470
 Qy 428 GAGATAATGATCAACAGCTAGCACTAGGCACTGTTCCACCATGGCTGAGCCCATGGAG 487
 Db 471 GAGATAATGATCAACAGCTAGCACTAGGCACTGTTCCACCATGGCTGAGCCCATGGAG 530
 Qy 488 ATGATTTCTCACCTGTTTATCAAGGAATGTTAGAGCAATTTCAACTCAGTTCCTTACA 547
 Db 531 ACCTATTTCTCACCTGTTTATCAAGGAATGTTAGAGCAATTTCAACTCAGTTCCTTACA 590
 Qy 548 ATTAATCTGATCAATCCAGTGTGTTACCAATGACAAACAAATTAATTTGGAGC 607
 Db 591 ATTAATCTGATCAATCCAGTGTGTTACCAATGACAAACAAATTAATTTGGAGC 650
 Qy 608 ATGGAGGATGCTGGTCAATGCAATTAATGCAAGTGAATTAATTAATCAAGATAAAC 667
 Db 651 ATGGAGGATGCTGGTCAATGCAATTAATGCAAGTGAATTAATTAATCAAGATAAAC 710
 Qy 668 CTAAGTTTGAAGTTCCATAAGGCTGGAATGTTTGGATTAATAACATATTATTGGGTTT 727
 Db 711 CTAAGTTTGAAGTTCCATAAGGCTGGAATGTTTGGATTAATAACATATTATTGGGTTT 770
 Qy 728 GTTTATATAAGTAGTGGATGTTGGTTTGGTACCAATTAATTAAGCTATGCTGTAATA 787
 Db 771 GTTTATATAAGTAGTGGATGTTGGTTTGGTACCAATTAATTAAGCTATGCTGTAATA 830
 Qy 788 TATAGGATTTATATTAACTATATCTGCATGCTTTTATATATAAAAAA 839
 Db 831 TATAGGATTTATATTAACTATATCTGCATGCTTTTATATATAAAAA 882

RESULT 5

US-10-412-699B-1631
 ; Sequence 1631, Application US/10412699B
 ; Publication No. US20040045049A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: Zhang, James
 ; APPLICANT: Fromm, Michael E.
 ; APPLICANT: Heard, Jacqueline E.
 ; APPLICANT: Riechmann, Jose Luis
 ; APPLICANT: Adam, Luc J.
 ; APPLICANT: Broun, Pierre E.
 ; APPLICANT: Pineda, Omaira
 ; APPLICANT: Reuber, T. Lynne
 ; APPLICANT: Keddle, James S.
 ; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Samaha, Raymond R.

; APPLICANT: Pilgrim, Marsha L.
 ; APPLICANT: Creelman, Robert A.
 ; APPLICANT: DuBell, Arnold N.
 ; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Kumimoto, Roderick
 ; APPLICANT: Sherman, Bradley K.
 ; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MBI-0048CIP
 ; CURRENT APPLICATION NUMBER: US/10/412,699B
 ; CURRENT FILING DATE: 2003-04-10
 ; PRIOR APPLICATION NUMBER: 09/394,519
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: 09/489,376
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: 09/506,720
 ; PRIOR FILING DATE: 2000-02-17
 ; PRIOR APPLICATION NUMBER: 09/533,030
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/533,392
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/533,029
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/532,591
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/533,648
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/713,994
 ; PRIOR FILING DATE: 2000-11-16
 ; PRIOR APPLICATION NUMBER: 09/819,142
 ; PRIOR FILING DATE: 2001-03-27
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 2011
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1631
 ; LENGTH: 968
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-412-699B-1631

Query Match 92.7%; Score 800.4; DB 18; Length 968;
 Best Local Similarity 97.4%; Pred. No. 9.1e-180; Indels 0; Gaps 0;
 Matches 810; Conservative 3; Mismatches 19;

Qy	8	CTCTATCACACACACAGTCAATGGATAAAAAACAACAGTGTGCAACGCTCTCAAGATCCT	67
Db	51	CTCTATCACACACACAGTCAATGGATAAAAAACAACGCTGTGCAACGCTCTCAAGATCCT	110
Qy	68	GAAGTGAGAAAGGGCCTTGACAAATGGAAGAGCTTGATCTTGATCAACTATATGCA	127
Db	111	GAAGTGAGAAAGGGCCTTGACAAATGGAAGAGCTTGATCTTGATCAACTATATGCA	170
Qy	128	AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACAGCTAAAGGA	187
Db	171	AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCTCAACAGCTAAAGGA	230
Qy	188	AAGAGTTCGGGCTAAGGTGGCTAAATTAATCTCGCTCTGATGTAGAGAGGAATATT	247
Db	231	AAGAGTTCGGGCTAAGGTGGCTAAATTAATCTCGCTCTGATGTAGAGAGGAATATT	290
Qy	248	ACACCCGAGGAGCAACTTTTGAATATGAGAGCTCAACGAAAGTGGGAAACAGGTGTCC	307
Db	291	ACACCCGAGGAGCAACTTTTGAATATGAGAGCTCAACGAAAGTGGGAAACAGGTGTCC	350
Qy	308	AAAATTGCGAAGCACTTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACA	367
Db	351	AAAATTGCGAAGCACTTACCTGGAGGACTGATATGAGATCAAGAACTATTGGAGGACA	410
Qy	368	AGGATCCAGAAGCAGATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATTTCT	427
Db	411	AGGATCCAGAAGCAGATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATTTCT	470
Qy	428	GAGATAATGATCAACAGCTAGCACTAGGCACTGTTCCACCATGGCTGAGCCCATGGAG	487

```
Db 471 GAGATAAATGATCCCAAGCTTGCACTAGCCATGTGTCCACCATGGCGCCCATGGAG 530
Qy 488 ATGATTTCTCCACCTGTATTCAAGGAATGTAGAGCCATTTTCAACTCAGTTCCTTACA 547
Db 531 ACCTATTCTCCACCTGTATTCAAGGAATGTAGAGCCATTTTCAACTCAGTTCCTTACA 590
Qy 548 ATTAATCCTGATCAATCCAGTGTGTACCAATGACAAACAACATTAATTAATTAATTA 607
Db 591 ATTAATCCTGATCAATCCAGTGTGTACCAATGACAAACAACATTAATTAATTAATTA 650
Qy 608 ATGAGGATAGCTGCTCAATCAATTAATCAAGCTGTATTAATTAATTAATTAATTAAT 667
Db 651 ATGAGGATAGCTGCTCAATCAATTAATCAAGCTGTATTAATTAATTAATTAATTAAT 710
Qy 668 CTAAGTTTGAAGTTTCCATAGGCTGGAATGTCTGTGATTAACATATTAATTAATTAAT 727
Db 711 CTAAGTTTGAAGTTTCCATAGGCTGGAATGTCTGTGATTAACATATTAATTAATTAAT 770
Qy 728 GTTTATATAGTAGTGGATGTTTGGTTTGGTGGTACCAATTAATTAATTAATTAATTAAT 787
Db 771 GTTTATATAGTAGTGGATGTTTGGTTTGGTGGTACCAATTAATTAATTAATTAATTAAT 830
Qy 788 TATACGAGATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839
Db 831 TATACGAGATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 882

RESULT 6
US-10-021-811-29
; Sequence 29, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-021-811-29

Query Match 70.7%; Score 610.2; DB 14; Length 988;
Best Local Similarity 92.0%; Pred. No. 1.4e-134;
Matches 724; Conservative 2; Mismatches 40; Indels 21; Gaps 7;

Qy 8 CTCTATCACACACAAAGTCAATGGATAAAAAACAACAGTGTAGAGCTTCAAGATCCT 67
Db 40 CTCTATCACACACAAACAAATGGATAAAAAACAACAGTGTAGAGCTTCAAGATCCT 99

Qy 68 GAAGTGAGAAAGGGCTTGGACAAATGGAGAGCACTTGATCTTGATGAACATATTTGCA 127
Db 100 GAAGTGAGAAAGGACCTTGGACCAATGGAGAGCACTTGATCTTGATGAACATATTTGCA 159

Qy 128 AATCATGGGAGAGTGTGTGGAACTCTTTGGCCAAAGCTGTGCTCAACAGTAAACGGA 187
Db 160 AATCATGGGAGAGTGTGTGGAACTCTTTGGCCAAAGCTGTGCTCAACAGTAAACGGA 219

Qy 188 AAGAGTTCGGGCTTAAGTGGCTAAATTAATCTCGTCTGATGTAGAGGGAATATT 247
Db 220 AAGAGTTCGGGCTTAAGTGGCTAAATTAATCTCGTCTGATGTAGAGGGAATATT 279

Qy 248 ACACCCGAGGACAACTTTTATGATGAGCTCCACCAAGTGGGGAACAGGTGTCC 307
Db 280 ACACCCGAGGACAACTTTTATGATGAGCTTCCACCAAGTGGGGAACAGGTGTCC 339
```

```
Qy 308 AAAATTGCCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAGAACTATTGGAGGACA 367
Db 340 AAAATTGCCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAGAACTATTGGAGGACC 399

Qy 368 AGGATCCAGAGGACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCT 427
Db 400 AGGATCCAGAGGACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCT 459

Qy 428 GAGATAAATGATCACCAGCTAGCAGTACCTAGCAGTGTTCACCATGGCTGAGCCCATGGAG 487
Db 460 GAGATAAATGATCACCAGCTAGCAGTACCTAGCAGTGTTCACCATGGCTGAGCCCATGGAG 519

Qy 488 ATGTATTCTCCACCTGTTATCAAGGAATGTAGAGCCATTTTCAACTCAGTTCCTTCCCT 544
Db 520 ACCTATTCTCCACCTTTTATCAAGGAATGTAGAGCCATTTTCTTCAATTCAGTTCCTCC 579

Qy 545 ACAATTAATCCTGATCAATCCAGTGTGTGACCAATGACAAACAACAGCATTTAACTAT 601
Db 580 ACAATTAATCCTGATCAATCCAGTGTGTGACCAATGACAAACAACAGCATTTAACTAT 639

Qy 602 TGGAGCATGGAGGATAGCTGCTCAATCAATTAATCAAGCTGTGATTAATTAATTAATTA 656
Db 640 TGGAGCATGGAGGATAGCTGCTCAATCAATTAATCAAGCTGTGATTAATTAATTAATTA 699

Qy 657 CAAGATAAAAACTAAGTGTG---AAGTTCAT--AAGGCTGGAATGTCTTGGATTAAA 711
Db 700 CAAGATAAAAACTAAGTGTG---AAGTTCAT--AAGGCTGGAATGTCTTGGATTAAA 759

Qy 712 ACA---TATTATTGGGTTGTTTATATATAGTAGTTGG--ATGTTGGTTTGGTACCAT 766
Db 760 ACATATTATTATTAGGTTTGTATATATAGTAGTTGGTATATGTTTGGTTTGGCTACCAT 819

Qy 767 TATTAGC 773
Db 820 TATTAGC 826

RESULT 7
US-10-659-869-29
; Sequence 29, Application US/10659869
; Publication No. US20040040057A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/659,869
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-659-869-29

Query Match 70.7%; Score 610.2; DB 18; Length 988;
Best Local Similarity 92.0%; Pred. No. 1.4e-134;
Matches 724; Conservative 2; Mismatches 40; Indels 21; Gaps 7;

Qy 8 CTCTATCACACACAAAGTCAATGGATAAAAAACAACAGTGTAGAGCTTCAAGATCCT 67
Db 40 CTCTATCACACACAAACAAATGGATAAAAAACAACAGTGTAGAGCTTCAAGATCCT 99

Qy 68 GAAGTGAGAAAGGGCTTGGACAAATGGAGAGCACTTGATCTTGATGAACATATTTGCA 127
```

```

Db 100 GAAGTGAGAAAGAGGACCTTGGACGATGAAGAAGACTTGTATCTTGTCAACTATATTGCA 159
Qy 128 AATCATGGGAGAGGTGTTTGGAACTCTTTGGCAAAAGCTGTGGTCTCAAAAGCTAAACGGA 187
Db 160 AATCATGGGAGAGGTGTTTGGAACTCTTTGGCAAAAGCTGTGGTCTCAAAAGCTAAACGGA 219
Qy 188 AAGAGTTCGCCGGCTAAGCTGAGTAAATTTACCTCCGTCCTGTATGTTAGAGAGGGAATATT 247
Db 220 AAGAGTTCGCCGGCTAAGCTGAGTAAATTTACCTCCGTCCTGTATGTTAGAGAGGGAATATT 279
Qy 248 ACACCCGAGGAAACAACTTTGATTTATGAGCTCCACGCAAAAGTGGGAAAACAGGTGTGCC 307
Db 280 ACACCCGAGGAAACAACTTTGATCATGAGCTTCAACGAAAAGTGGGAAAACAGGTGTGCC 339
Qy 308 AAAATTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACA 367
Db 340 AAAATTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACC 399
Qy 368 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTCT 427
Db 400 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAAATCAGCAATAACTCT 459
Qy 428 GAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 487
Db 460 GAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCTACCATGGCTGAAACCCCATGGAG 519
Qy 488 ATGATTCTCCACCCCTGTTTATCAAGGAATGTTAGAGCCATT---TTCAACTCAGTTCCTT 544
Db 520 ACCTATTCTCCACCCCTTTTATCAAGGAATGTTAGAGCCATTTTCTCAATTCAGTTCCTCC 579
Qy 545 ACAATTAATCTCGATCAATCCAGTTGTTGTACCAATGACAAACAA---CATTAACATPAT 601
Db 580 ACAATTAATCTCGATCAATCCAGTTGTTGTACCAATGACAAACAAACAGCATTAACTAT 639
Qy 602 TGGAGCTGAGGATAGCTGTGTCATGCAATTAAGTCAAGCGTGATTAATAT---TAT 656
Db 640 TGGAGCTGAGGATATCTGTCATGCAATTAAGTCAAGCGTGATTAATAT---TAT 699
Qy 657 CAAGATAAAACCTAAGTTGTTG---AAGTTTCCAT---AAGCTGGAATGCTVTGGATTAAA 711
Db 700 CAAGATAAAACCTAATTTCTGTATTAAGTTCCATAAACAACACTGGAATGCTCTGGCTTAAA 759
Qy 712 ACA---TATTATTGGTGTGTTTATATAGTAGTTGG---ATGTTTGGTTTTCGGTACCAT 766
Db 760 ACATATTATTATTAGTTTGTGTTTATATAGTAGTTGGATATGTTTGGTTTTCGGTACCAT 819
Qy 767 TATTAGC 773
Db 820 TATTAGC 826

```

```

RESULT 8
US-10-424-599-85949
; Sequence 85949, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85949
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48625C.1
US-10-424-599-85949

```

```

Query Match 54.6%; Score 471; DB 18; Length 1097;
Best Local Similarity 82.8%; Pred. No. 1.7e-101;
Matches 640; Conservative 2; Mismatches 52; Indels 79; Gaps 6;

Qy 8 CTCTATCACACACAAAGTCAATGGATATAAATAAACAACAGTGTAAAGAGCTCTCAAGATCCT 67
Db 49 CTCTATCACACACAAATAAATAAATAAACAACAGTGTAAAGAGCTCTCAAGATCCT 108
Qy 68 GAAGTGAGAAAGGCGCTTCGACAAATGGAAGAGACTTGTCTTGTGATGAACATATTATGCA 127
Db 109 GAAGTGAGAAAGGCGCTTCGACAAATGGAAGAGACTTGTCTTGTGATGAACATATTATGCA 168
Qy 128 AATCATGGGAGAGGTGTTTGGAACTCTTTGGCCAAAAGCTGTCTCTCAAAACGTAACGGA 187
Db 169 AATCATGGGAGAGGTGTTTGGAACTCTTTGGCCAAAAGCTGTCTCTCAAAACGTAACGGA 228
Qy 188 AAGAGTTCGCCGGCTAAGCTGAGTAAATTTACCTCCGTCCTGTATGTTAGAGAGGGAATATT 247
Db 229 AAGAGTTCGCCGGCTAAGCTGAGTAAATTTACCTCCGTCCTGTATGTTAGAGAGGGAATATT 288
Qy 248 ACACCCGAGGAAACAACTTTTGTATGAGCTCCACGCAAAAGTGGGAAAACAGGTGTGCC 307
Db 289 ACACCCGAGGAAACAACTTTTGTATCATGAGCTCCACGCAAAAGTGGGAAAACAGGTGTGCC 348
Qy 308 AAAATTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACA 367
Db 349 AAAATTGCCAAGCATCTACCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGGACA 408
Qy 368 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTCT 427
Db 409 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAAATA--- 455
Qy 428 GAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAG 487
Db 456 -----GTGTCACCATGGCGCAACCATGGAA 482
Qy 488 ATGTATTCTCCACCTGTTTATCAAGGAATGTTAGAGCAATTTTCAACTCAGTTCCTCTACA 547
Db 483 ACCTATTCTCCACCCAGTTATCAAGGAATGTTAGATCCATTTTCAATTCAGTTCCTCAACA 542
Qy 548 ATTAATCTGATCAATCCAGTTGTTGTACCAATGACACACACACATTAATATTGGAGC 607
Db 543 ---AATCCTCATCATTTAGTTGTTGTACCAATGACGACGACAAACAACATATTGGAGC 599
Qy 608 ATGAGGATAGCTGCTCAATG-----CAATTAAGTGTGAAAGTTCATTAAGGCTGG-- 650
Db 600 ATGAGGATAGCTGCTCAATGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCAATTAAG 659
Qy 651 -----TATTATCAAGATAAAGCTTAAGTGTGAAAGTTCATTAAGGCTGG--AATGTC 700
Db 660 ATAAATATAAATAATATATAAACCCTAAGTGTCTAAGTTCATTAAGTTCATTAAGTGTCT 719
Qy 701 TTTGGATTAAACATATTATTGGTGTGTTTATATAGTAGTGTGATGTTGG 753
Db 720 TCTGCTTTAAACATGTTA---GGTTGTTTATACAAGTAGTGTGATGTTGG 769

```

```

RESULT 9
US-10-374-780A-1538
; Sequence 1538, Application US/10374780A
; Publication No. US2004001927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James

```

```

; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omita
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR FILING DATE: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1538
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1322
US-10-374-780A-1538

Query Match          49.0%; Score 423; DB 17; Length 969;
Best Local Similarity 81.5%; Pred. No. 4.2e-90;
Matches 573; Conservative 2; Mismatches 92; Indels 36; Gaps 6;

Qy 81 GGCCTTGACATGGAAGAAGACTTGATCTGTGAGTAATATTTGCAAAATCATGGGGAAG 140
Db 44 GGGCCAGGACAATGGAAGAAGACTTAATCTTTGATCACCTATATTTGCCAATCACGGGGAAG 103
Qy 141 GTGTTTGGAACTTTTGGCCAAAGCTGCTGTCTCAAAAGTGAACGGAAAGATTGCCGGC 200
Db 104 GGGTTTGGAACTTTTGGCCAAAGCTGCTGTGACTTAAACGTACCGGAAAGATTGCCGGC 163
Qy 201 TAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAAGAGGGAATATTACACCCGAGGAAC 260
Db 164 TTCGGTAGCTAAACTACCTAGCTCCTGATGTTAGAAGAGGGAATATTACACCCGAGGAAC 223
Qy 261 AACTTTTGTATATGAGAGCTCCACCAAGTGGGGAACAGGTGTCACAAATTCGCCAAGC 320
Db 224 AGCTTTTGTATGATGGAACCTTCATCAAAAGTGGAGCAACAGGTGTCACAAATTCGCCAAGC 283
Qy 321 ATCTACCTGGAAGACTGATTAATCAGATCAAGACTATTGGAGGACAGGATCCAGAGC 380
Db 284 ATCTACCCGGAAGACTGATTAATGAGATTAAGAACTACTGGAGGACAGGATCCAGAGC 343
Qy 381 ACATCAAGCAAGCT---GAGAACTTTTCAGCAACAGAGTAGTAATAATTCTGAGATAAATG 437
Db 344 ACCTCAAGCAAGCTTCCAGCAGCTTCCAGCAACAGAGTAGTAATTCTGAGATAATTATC 403
Qy 438 ATCAACCAAGCTAGCACTAGCCATGTTTCCACCATGCTGAGCCCATGAGATGATTCTC 497
Db 404 ATCCCAAGCTTGCACTAGCAAGCTGTCACCATGCTGAGCCCATGAGATGATTCTC 463
Qy 498 CACCTGTATCAAGGAATGTTAGAGCACTTTTCAACTCAGTTCCCTCAATTAATCCTG 557
Db 464 CACCAGTTATCAAGGAATGTTAGATCCATTTTCAATTCAGTTCCCAACA---AATCCTC 520

; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omita
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1632
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Glycine max
US-10-412-699B-1632

; Sequence 1632, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omita
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1632
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Glycine max
US-10-412-699B-1632

558 ATCAATCCAGTTGTTGTACCAATGACAAACAACAACATTAATACTATTGGAGCATGGAGATA 617
Db 521 ATCATCTAGTTGTTGTACCAATGACGACGACAAACAACATTAATTGGAGCATGGAGATA 580
Qy 618 GCTGGTCAATG-----CAATTACTGACGGTGATTAA-----TA 652
Db 581 TCTGGTCAATGCAATAGCAATAGCAATTAAGTCTAAGTTCCATAAATAAGCTGTAGTCTCTGGCTTAA 640
Qy 653 TTATCAAGATAAAAAACCTTAAGTTTGAAGTTCCATAAGCGCTGG--AATGTCTCTTGGATTAA 710
Db 641 ATAATATATATAAACCTTAAGTCTAAGTTCCATAAATAAGCTGTAGTCTCTGGCTTAA 700
Qy 711 AACATATATTGGGTTTGGTTTATATATAGTAGTGGATGTTGG 753
Db 701 AACATGTTA---GGTTTGTATTATACAAGTAGTGGATGTTGG 740

RESULT 10
US-10-412-699B-1632
; Sequence 1632, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omita
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1632
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Glycine max
US-10-412-699B-1632
```

Query Match 49.0%; Score 423; DB 18; Length 969;
Best Local Similarity 81.5%; Pred. No. 4.2e-90;
Matches 573; Conservative 2; Mismatches 92; Indels 36; Gaps 6;
QY 81 GGCCTTGACAAATGGAAGAAGACTTGATCTTGATGAACATATATTGCAAAATCATGGGGAAG 140
DB 44 GGGCCAGGACAAATGGAAGAAGACTTAATCTTTGATCACCTATATTGCCAATCACGGGGAAG 103
QY 141 GTGTTTGAACCTTTTGCCCAAGCTGCTGCTCAAAAGCTAAACGGTAACGGGAAGAGTTGCCGGC 200
DB 104 GGGTTTGAACCTTTTGCCCAAGCTGCTGCTCAAAAGCTTAACCGTACCGGGAAGAGTTGCCGGC 163
QY 201 TAAGTGCTCTAAATTAACCTCTCGTCTGATGTTAGAAAGAGGAATATTACACCCGAGGAAC 260
DB 164 TTGGGTAGCTAAATTAACCTCTGATGTTAGAAAGAGGAATATTACACCCGAGGAAC 223
QY 261 AACCTTTTGATATGGAGCTCCACGCAAGTGGGGAACACAGTGGTCCAAAATTTGCCAAGC 320
DB 224 AGCTTTTGATCATGGAACCTTCATGCAAAAGTGGAGCAACAGGTGGTCCAAAATTTGCCAAGC 283
QY 321 ATCTACCTGGGAAGACTGATTAATGAGATCAAGAACTATTGGAGGACAGAGATCCAGAGC 380
DB 284 ATCTACCCGGAAGGACTGATTAATGAGATTAAGAACTACTGGAGGACAAAGGATCCAGAGC 343
QY 381 ACATCAAGCAAGCT--GAGAACTTTTCAGCAACAGAGTAGTAATAATTTCTGAGATAAATG 437
DB 344 ACCTCAAGCAAGCTTCAGAGCTTCAGCAACAGAGTAGTAATAATTTCTGAGATAAATG 403
QY 438 ATCAACCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGATTTCTC 497
DB 404 ATCCCAAGCTTGCACTAGCACTAGCAAGTGTCCACCATGGCGCAGCCCATGAAACCTATTCTC 463
QY 498 CACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTTCAATTAATCTCTG 557
DB 464 CACCAGTTATCAAGGAATGTTAGATCCATTTTCAATTTCAATTTCAATTTCAATTTCAATTT 520
QY 558 ATCAATCAGTGTGTTGTTACCAATGACAAACAACTTAATTTATTTGGAGCATGGAGGATA 617
DB 521 ATCATCTAGTGTGTTGTTACCAATGACGAGCAACAACTATTTGGAGCATGGAGGATA 580
QY 618 GCTGCTCAATG-----CAATTAAGTGAACGGTGAATTA-----TA 652
DB 581 TCTGCTCAATGCAATTAGCCAAATTAAGTGAACGGGAATTAACCTTAAACATAAATAATAA 640
QY 653 TTATCAAGATAAAACCTAAGTGTGAGTTCATTAAGCTG--AATGCTCTTGTGATTA 710
DB 641 ATAATATATATAAACCTAAGTGTCTAAGTTCATTAATTAAGCTGTGATCTCTGCTTAA 700
QY 711 AACATATTATTGGGTTGTTTATATAAGTAGTTGGATGTTGG 753
DB 701 AACATGTTA--GGTTTGTATACAAAGTAGTTGGATGTTGG 740

RESULT 11
US-10-424-599-126212
; Sequence 126212, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 126212
; LENGTH: 1094
; TYPE: DNA
; ORGANISM: Glycine max

FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84978C.1
US-10-424-599-126212

Query Match 47.5%; Score 409.6; DB 18; Length 1094;
Best Local Similarity 70.6%; Pred. No. 6.7e-87;
Matches 597; Conservative 2; Mismatches 221; Indels 26; Gaps 3;
QY 7 GCTCTATCACACACAAAGTCAATGGATTAATAAACAACAGTGAAGAGCTCTCAAGATCC 66
DB 10 GCTCGAATGACACAAAGTAAATGGACAAATAAACCAGTCACTCATCTCATGATCC 69
QY 67 TGAAGTGAAGAAAGGGCTTGGACAAATGGAAGAAGACTTGATCTTTGATGAACATATATTGC 126
DB 70 TGAAGTGAAGAAAGGGACCATGGACCATGGAAGAAGACTTGATCTTTGATGAACATATATTGC 129
QY 127 AAATCATGGGAAAGCTGTTGGAACTCTTTGGCCAAAGCTGCTGCTCAACCTTAACGG 186
DB 130 AAATCACGGTGAAGGTGTTGGAACTCTTATAGCCAAAGCTTCTGCTCTCAAAACGAACGG 189
QY 187 AAAGAGTTGCCGGCTTAAGGTGCTAAATTAACCTCCGCTCTGATGTTAGAAAGAGGAATAT 246
DB 190 AAAGAGTTGTCGACTCCGTTGGCTAAACTACTCTGCTGCTGATGTTAGAAAGGAACAT 249
QY 247 TACACCCGAGGAACAACTTTTATTTATGGAGCTCCACGCAAAAGTGGGGAACAGGTGTC 306
DB 250 TACACCCGAGGAACAGCTTTTATGATCATAGAACTTTCATGCAAAAGTGGGCAATAGGTGTC 309
QY 307 CAAAATTTGCCAGCATCTACCTGGAAGGACTGATTAATGAGATCAAGAACTATTGGAGGAC 366
DB 310 CAAAATTTGCAAGGATCTTCCAGGAAGAACTGCAATGAGATTAAGAACTTCTGGAGAAC 369
QY 367 AAGGATCCAGAAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTC 426
DB 370 AAGGATCCAAAGGACATTAAGCAAGCTGAGACTTTCACAAACAACTGATTAATCAGAGAA 429
QY 427 TGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCAGTGGCTGAGCCCATGGA 486
DB 430 TAATGATCATCAAGCAAGCACTAGTACTAGCAAAAGTGTCCACCATGGCACATCCAAATGA 489
QY 487 GATGATTTCTCCAGCTGTTATCAAGGAATGTTAGAGCCATTTTCACTCAGTTCCCTAC 546
DB 490 GACTTTCTCTCCAGCTCATACCAAGGAACCTTTTGAGCCATTTCAACCTCAATTTCCCTAC 549
QY 547 AATTAATCCTCATCAATCCAGTGTGTTGACCAATGACAAACAAACAACTAATTAATTTGGAG 606
DB 550 AATCA--CATCAATCAAGTGTGTTGACCAACCAACCAACCAACCAACCACTATTGGAG 606
QY 607 CATGAGGATAGCTG---GTCAATGCAATTAAGTGAACGGTGAATTA----- 649
DB 607 CATGAGGATATCTGGTCTGCTATGCAATTAAGTGAATTAAGGATTAAGCTATATATGC 666
QY 650 -----ATATTTATCAAGATAAAACCTAGTGTGAGTTTCCATAAGGCTGGAATGCTTCT 703
DB 667 ATGCTATATAAATCATATATATATATATATAAACCCTAAGCTCTTGTAGAGTGTGTTCT 726
QY 704 GGATTTAAACATATTTATTTGGGTTGTTTATATAAGTAGTTGGATGTTTGGTTTCCGTAC 763
DB 727 AGGCTTAATAACATCATTTAGTCTGTTTATATAGTAGTCTTAAGTGTGTTGTTGTTAATG 786
QY 764 CATTATTAGCTATGCTGTTAATATATACGAGATVTTTATTAATAACTATATCTGCAATGCT 823
DB 787 CATGATGTGAGTTAAGAAATTAATTTAGTTATGGATTGAATAATATATAGTAACCTTATATCT 846
QY 824 TTATAT 829
DB 847 ACATGT 852

RESULT 12
US-10-425-114-10688
; Sequence 10688, Application US/10425114
; Publication No. US20040034888A1


```
Qy 498 CACCTGTTATCAAGGAATGTTAGAGCAATTTTCAACTCAGTTCCTTACAAATTAATCCTG 557
Db 529 CACCTCATATCAAGGAATTTTGGAGCAATTTTCAACTCAGTTCCTTACAAATTAATCCTG 585
Qy 558 ATCAATCAGTTGTTGACCAATGACAAACAACATTAATTAATTTGGAGCATGGAGATA 617
Db 586 ATCAATCAAGTTGTTGACCAATGACAAACAACATTAATTAATTTGGAGCATGGAGATA 645
Qy 618 GCTG---GTCATCAATTAATGACGAGTGAATTA-----ATATT 654
Db 646 TCTGGTCTGTCATGCAATTAATGACGAGTGAATTA-----ATATT 705
Qy 655 ATCAAGTAACAACCTAAGTTTGAAGTTCCATAAGGCTGGATCTCTTTGGAATTAACA 714
Db 706 AATCATATATGATGATATATAATAAACCCTTAAGCTCTTTGAGAGTGTGTCAGGCTTAATA 765
Qy 715 TATTATTGGGTTGTTTATATAAGTAGTTGGATGTTGGTTTTCGCTACCAATTAATAGCT 774
Db 766 CATCATTTAGGTCGTTTATATAGTAGTCTTAAGTTTGGTGTGTTGTAATGATGATGTGAG 825
Qy 775 ATGTGCTGTAATATACGAGATTTTATATTAATACTATATCTGCATGCTTTTATAT 829
Db 826 TTAAGAATTAATTTAGTTAGTTGAATTAATAATATAGTAACCTATATATACATCT 880

RESULT 15
US-10-425-114-10695
; Sequence 10695, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10695
; LENGTH: 877
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700960664_FLI
US-10-425-114-10695

Query Match 46.3%; Score 399.8; DB 18; Length 877;
Best Local Similarity 70.7%; Pred. No. 1.3e-84;
Matches 583; Conservative 2; Mismatches 214; Indels 26; Gaps 3;

Qy 28 AATGGATAAATAAACAAGAGTGAAGAGTCTCAAGATCTCAAGTGAAGAAAGGCGCTTG 87
Db 3 AATGGCAAAAAAACCATGCAACTCATCATCTCATGATCTCTGAAGTGAAGAAAGGCAACATG 62
Qy 88 GACAATGGAAGAGAGCTTGTGATGAACTATATTTGCAAAATCATGGGAGGTGTTG 147
Db 63 GACCATGGAAGAGAGCTTGTGATGAACTATATTTGCAAAATCATGGGAGGTGTTG 122
Qy 148 GAACCTTTTGCCAAAGCTGCTGCTCAACAGTGAAGGAAAGAGTTGCGGCTTAAGGTG 207
Db 123 GAACCTTTTAGCAAAAGCTTCTGCTCAACAGTGAAGGAAAGAGTTGCGACTCCGTTG 182
Qy 208 GCTAAATTAATCCTCGTCTGATGTTAGAGGGAATATTACCCGAGGAAACAATTTTT 267
Db 183 GCTAAATTAATCCTCGTCTGATGTTAGAGGGAATATTACCCGAGGAAACAATTTTT 242
Qy 268 GATTATGAGCTCCACCAAGTGGGAAACAGGTGCTCAAAATTTGCCAAGCATCTACC 327
Db 243 GATCATAGAACTTCATGCAAAAGTGGGCAATAGGTGCTCAAAATTTGCCAAGCATCTTC 302
```

```
Qy 328 TGAAGGAGCTGATAATGAGATCAAGAACTATTGGAGGACCAAGGATCCAGAGCAATCAA 387
Db 303 AGGAAGAACTGACAAATGAGATTAAAGAACTTCTGGAGAAACAAGGATCCAAAAGCACATTAA 362
Qy 388 GCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCTCGAGATAAATGATCACCAGC 447
Db 363 GCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCTCGAGATAAATGATCACCAGC 422
Qy 448 TAGCACTAGCACTGTTTCCACCATGGCTGAGCCCATGGAGATGTTATCTCCACCTGTTA 507
Db 423 TAGTACTAGCAAGAGTGTCCACCATGGCACATCCAAATGAGACTTTCTCTCCACCTCAT 482
Qy 508 TCAAGGAATGTTAGAGCAATTTTCAACTCAGTTCCTTACAAATTAATCTCGATCAATCCAG 567
Db 483 CCAAGGAATTTTGGAGCAATTTCAACCTCAATTCCTTACAATCA--CTGATCAATCAAG 539
Qy 568 TTGTTGTTACCAATGACAACAACAACATTAACATTTGAGCATGGAGGATAGCTG---GTC 624
Db 540 TTGTTGTTACCAACCAACCAACCAACCAACCAACTATTGGAGCATCGAGGATATCTGGTCGTC 599
Qy 625 AATGCAATTACTGAAACGGTGATTAA-----ATATTATCAAGATAA 664
Db 600 TATGCAATTACTCAATGGAGATTAAACCTAGCTATATATGCTATATAATCATATAT 659
Qy 665 AACCTAAGTTTGAAGTTCCATAAGGCTGGAATGTCTTGTGATTTAAACATATTAATTTGGG 724
Db 660 ATGATGATATATAAACCCTAAGCTCTTGTAGAGTGTGTTCAAGGCTTAATAACATCATTAGG 719
Qy 725 TTTGTTTATATAAGTAGTTGGATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 784
Db 720 TCTGTTTATATAGTAGTCTTAAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 779
Qy 785 ATATATACGAGATVYTTATATAATAAATACTATATCTGCATGCTTTTATAT 829
Db 780 ATTGATTGAGATTGAATATATATAGTAACCTTATATACTATACATGT 824
```

Search completed: September 28, 2005, 13:31:28
Job time : 744 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 07:25:22 ; Search time 3558 Seconds
(without alignments)
9232.565 Million cell updates/sec

Title: US-10-659-869A-35
Perfect score: 863
Sequence: 1 gcacgagctctatcacac.....aaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_hic:*
 - 4: gb_est3:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_gss1:*
 - 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	65.1	562	4	BM732121 sal173b11
2	506.4	58.7	586	2	BE820766 GM700012A
3	491.4	56.9	782	2	BE658316 GM700005B
4	482.6	53.6	501	4	BM527606 sal63906
5	444.4	51.5	450	2	BS058947 sn23b01.y
6	442.8	51.3	523	1	A1930997 sb4sh07.y
7	425.4	49.3	431	2	AW459279 sh22h08.y
8	424.4	49.2	552	4	BM527774 sal65903
9	401.2	46.5	568	4	BM527664 sal64d10
10	400.6	46.4	544	4	BM527508 sal62f08
11	369.8	42.9	409	2	BE804790 ss45f11.y
12	369.6	42.8	560	4	BM732539 sal78e08
13	368.6	42.7	635	5	BQ146360 NF047E06F
14	363.6	42.1	475	2	BR057370 sn02c04.y
15	347	40.2	542	4	BM528383 sal57f09
16	332	38.5	522	5	BP032853 BP032853
17	330.8	38.3	425	2	AW423958 sh58e05.y
18	307.8	35.7	682	5	BQ146831 NF028G06F
19	306.2	35.5	688	4	B1272897 NF091A12F
20	304.8	35.3	669	4	BG457971 NF037A10P
21	304.6	35.3	395	2	BR805071 ss37a07.y
22	294	34.1	470	2	AW156542 se28d11.y
23	291.2	33.7	447	4	B1273011 NF097F02F
24	288.4	33.4	558	2	BE324639 NF024C04P

c	25	284	32.9	766	2	BE659054
	26	282	32.7	659	7	CV130213
	27	282	32.7	663	7	CK468561
	28	279.4	32.4	715	2	BE658737
c	29	279	32.4	567	4	BM732162
	30	277	32.1	638	7	CV466307
	31	276	32.0	570	7	CV298677
	32	275	31.9	592	2	AW928296
	33	273.8	31.7	452	7	CN495741
	34	271.6	31.5	500	1	AI897784
	35	271.6	31.5	519	1	AI897681
	36	270.2	31.3	567	7	CO414782
	37	265.6	30.8	397	2	AW432364
	38	265.4	30.8	681	1	AI486576
	39	261	30.2	805	7	CO113188
	40	258.8	30.0	503	5	BQ104458
	41	254.4	29.5	616	7	CF349392
	42	252.2	29.2	761	1	AJ791605
	43	249.2	28.9	658	1	AJ790280
	44	246	28.5	502	5	BQ106505
	45	243.6	28.2	582	4	BI977753

ALIGNMENTS

RESULT 1
BM732121 562 bp mRNA linear EST 06-JUL-2004
LOCUS sal173b11.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl061-4678 5' similar to TR:Q39028 Q39028 ATMB2.1; mRNA
sequence.

ACCESSION BM732121
VERSION BM732121.1 GI:19053454
KEYWORDS EST
SOURCE Glycine max (soybean)
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 562)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCamr,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES
Location/Qualifiers
1..562
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Raiden"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-4678"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

```

Query Match      65.1%; Score 562; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 4e-116;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CTATCACACACACAGTCAATGGATATAAAAAACACAGGTGAAGCGTCTCAAGATCCCTGA 69
DB 1 CTATCACACACACAGTCAATGGATATAAAAAACACAGGTGAAGCGTCTCAAGATCCCTGA 60

QY 70 AGTGAGAAAAGGCGCTTCGACAAATGGAAGAGACTTGATCTTGATGAAGTATATGCAAA 129
DB 61 AGTGAGAAAAGGCGCTTCGACAAATGGAAGAGACTTGATCTTGATGAAGTATATGCAAA 120

QY 130 TCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAACTGAACGAA 189
DB 121 TCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAACTGAACGAA 180

QY 190 GAGTTGCGCGCTTAAGTGGCTAAATTAACCTCGTCTGATGTTAGAAGAGGAATATTAC 249
DB 181 GAGTTGCGCGCTTAAGTGGCTAAATTAACCTCGTCTGATGTTAGAAGAGGAATATTAC 240

QY 250 ACCCGAGGAACAACCTTTTGAATATGAGAGCTCCACGAAAGTGGGAAACAGGTGGTCCAA 309
DB 241 ACCCGAGGAACAACCTTTTGAATATGAGAGCTCCACGAAAGTGGGAAACAGGTGGTCCAA 300

QY 310 AATTGCCAAGCATCTACCTGGNAGAGCTGATAATGAGATCAAGACTATTGGAGGACAAG 369
DB 301 AATTGCCAAGCATCTACCTGGNAGAGCTGATAATGAGATCAAGACTATTGGAGGACAAG 360

QY 370 GATCCAGAGACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTATAATTCGA 429
DB 361 GATCCAGAGACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTATAATTCGA 420

QY 430 GATAAATGATCACCAGCTAGCATTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGAT 489
DB 421 GATAAATGATCACCAGCTAGCATTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGAT 480

QY 490 GTATTCTCGACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTTACAAAT 549
DB 481 GTATTCTCGACCTGTTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTTACAAAT 540

QY 550 TAATCCTGATCAATCCAGTTGT 571
DB 541 TAATCCTGATCAATCCAGTTGT 562

```

RESULT 2

```

BE820766/c
LOCUS      BE820766 586 bp mRNA linear EST 24-MAY-2001
DEFINITION GM700012A20H11 Gm-r1070 Glycine max cDNA clone Gm-r1070-4966 3',
            mRNA sequence.
ACCESSION  BE820766
VERSION    BE820766.1 GI:10253000
KEYWORDS   EST.
SOURCE     Glycine max (soybean)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 586)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corvelli, V.,
Expelting, J., Raph, C., Shoop, E., Padias, J., Liu, L. and Lewin, H.
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AW459279 corresponding to Gm-cl016-5560 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTATTTTATTTT(A/C/G)-3'.

FEATURES

source

```

1..586
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone_lib="Gm-r1070-4966"
/clone_lib="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, reracked
set of 9,216 clones selected from cDNA libraries from
various tissues and stages of development of soybean that
represent 2,639 sequences from immature cotyledons, 1,770
from immature seed coats, 3,938 from flowers, and 869
from young pods. The 5' ESTs of the source clones from
the different libraries was used to select singletons, or
a representative of each contig, which were reracked to
form library Gm-r1070. The cDNA clones of the reracked
Gm-r1070 library were then sequenced at the 3' end. The
contig analysis to select unique genes was performed by
the laboratory of Ernest Retzel, Center for Computational
Genomics and Bioinformatics, University of Minnesota,
http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
. Reracking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and 3' sequencing by the
Keck Center for Comparative and Functional Genomics,
University of Illinois,
http://www.life.uiuc.edu/biotech/keck.html. Note: The
corresponding 5' EST from each clone in the Gm-r1070
library is listed in the 'OTHER EST' field. The detailed
information on the source library for each clone can also
be obtained by referring to the Genome Systems clone ID of
the original cDNA library that is also listed under
'OTHER EST'."
```

ORIGIN

```

Query Match      58.7%; Score 506.4; DB 2; Length 586;
Best Local Similarity 98.6%; Pred. No. 1.4e-103;
Matches 507; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 326 CCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGCAAGGATCCAGAAGCACATC 385
DB 583 CCTGGAAGGACTGATATGAGATCAAGAACTATTGGAGCAAGGATCCAGAAGCACATC 524

QY 386 AAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTATAATTTCTGAGATAAATGATCACC 445
DB 523 AAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTATAATTTCTGAGATAAATGATCACC 464

QY 446 GCTAGCACTAGCCATGTTTCCACCATGCTGAGCCCATGGAGATGTTATTTCCACCTGT 505
DB 463 GCTAGCACTAGCCATGTTTCCACCATGCTGAGCCCATGGAGATGTTATTTCCACCTGT 404

QY 506 TATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTTCCCTACAAATTAATTCCTGATCAATCC 565

```

403 TATCAAGGAATGTAGAGCCATTTTCAACTCAGTTCCTACAAATTAATCTGTATCAATCC 344
QY
566 AGTTGTTGTACCAATGACAAACAACATTAATATTGGAGCATGGAGTAGCTGTCA 625
DB
343 AGTTGTTGTACCAATGACAAACAACATTAATATTGGAGCATGGAGTAGCTGTCA 284
QY
626 ATGCAATTAATCAAGCGTGAATTAATTAATCAAGATATAAACAATTAAGTTTGAAGTTCCA 685
DB
283 ATGCAATTAATCAAGCGTGAATTAATTAATCAAGATATAAACAATTAAGTTTGAAGTTCCA 224
QY
686 TAAGGCTGGATCTCTGTGATTAACAAATTAATTTGGGTTTCTTTATATAAGTACTGG 745
DB
223 TAAGGCTGGATCTCTGTGATTAACAAATTAATTTGGGTTTCTTTATATAAGTACTGG 164
QY
746 ATGTTTGGTTTGGTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 805
DB
163 ATGTTTGGTTTGGTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 104
QY
806 AACTATATCTGATGCTTTTATATAAATAA 839
DB
103 AACTATATCTGATGCTTTTATATAAATAA 70

RESULT 3
BE58316/c
LOCUS
DEFINITION
BE58316
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Other ESTs: A1930997 corresponding to Gm-cl015-278 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomeystems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTATTTTATTTT(A/C/G)-3'.
Location/Qualifiers
1. 782
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1070-1759"
/clone_lib="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, reracked set of 9,216 clones selected from cDNA libraries from various tissues and stages of development of soybean that represent 2,639 sequences from immature cotyledons, 1,770 from immature seed coats, 3,938 from flowers, and 869 from young pods. The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to

form library Gm-r1070. The cDNA clones of the reracked Gm-r1070 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html. Reracking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 56.9%; Score 491.4; DB 2; Length 782;
Best Local Similarity 89.3%; Pred. No. 3.4e-100;
Matches 612; Conservative 2; Mismatches 49; Indels 22; Gaps 8;
QY 111 TGATGAATATATTCGAATCATGGGAAGTCTTTGGAACTCTTTGGCCAAAGCTCTG 170
DB 782 TGATGAATATATTCGAATCATNNNNNGGTGTGTGNANNNNTTGGCCAAAGCTCTN 723
QY 171 GTCTCAAACTGAACGGAAGAGTTGCGGCTAAGGTGGCTAAATTTACTCCGCTCTGATG 230
DB 722 NTCTCAACGCTACCGGAAGAGTTGCGGCTAAGGTGGCTAACTACTCCGCTCTGATG 663
QY 231 TTAGAAGAGGGAATATATACACCGGAAGAACTTTTGTATTTAGAGCTCAGCAAGT 290
DB 662 TTAGAAGAGGGAATATNNNACCAGGAACAATTTTGTATCATGAGCTTTCAGCAAGT 603
QY 291 GGGGAACAGGTGGTCCAAATTTCCAAAGCATCTTCTGGAAGGACTGATAAGATGATCA 350
DB 602 GGGGAACAGGTGGTCCAAATTTGCCAAGCATCTTCTGTTGAGACAGATAAGATGATCA 543
QY 351 AGAATCTATTGGAGGACCAAGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAAC 410
DB 542 AGAATCTATTGGAGGACCAAGGATCCAGAGCACATCAAGCAAGCTGAGAACTTTTCAGCAAC 483
QY 411 AGAGTAGTAATAATCTGAGATAAATGATCAACAAGTACACATAGCCATGTTTCCACCA 470
DB 482 AAATCAGCAATAACTCTGAGATAAATGATCAACAAGTACACATAGCCATGTTTCTACCA 423
QY 471 TGCTGAGCCCATGGAGATGTTTCTCCACCTGTTATCAAGGAATCTTAGAGCCATT - 528
DB 422 TGCTGAGCCCATGGAGACCTATTTCTCCACCTTTTATCAAGGAATCTTAGAGCCATT 363
QY 529 -TTCAACTCAGTTCCCTTACAAATTAATCTGATCAATCCAGTGTGTGTACCAATGACAACA 587
DB 362 CTTCAATTCAGTTCCCAACAATTAATCTGATCAATCCAGTGTGTGTACCAATGACAACA 303
QY 588 ACAAA---CAATTAATCTGAGGATGAGGATAGCTGGTCAATG-CAATTAATCTGAAAGCT 643
DB 302 ACAACAGCAATTAATCTGAGGATGAGGATATCTGGTCAATGATAGTACTGAAAGCT 243
QY 644 GATTAATAT-----TATCAAGATAAACCCTAAGTTTG---RAGTTCCAT--RAGGCTG 693
DB 242 GATTAATATTTGATATATCAAGATAAACCCTAAGTTTCTGTATAGTTCCTAAACACTG 183
QY 694 GAATGCTCTGGATTAAACA---TATTATTTGGTTTGTATATATAAGTAGTATGATG 748
DB 182 GAATGCTCTGGCTTAACAATATATATATAGTGTGTTTATATAAGTAGTATGATG 123
QY 749 TTTGGTTTTCGTAACATTTATAGC 773
DB 122 TTTGGTTTTCGTAACATTTATAGC 98

RESULT 4

```

BM527606          501 bp      mRNA      linear      EST 06-JUL-2004
LOCUS             sal63906.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION        Gm-cl061-3780 5' similar to TR:Q9028 Q39028 ATMVB2. ;, mRNA
sequence.
ACCESSION         BM527606
VERSION           BM527606.1 GI:18733434
KEYWORDS          EST.
SOURCE            Glycine max (soybean)
ORGANISM          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                  Glycine.
REFERENCE         1 (bases 1 to 501)
AUTHORS           Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
                  Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
                  Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
                  Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
                  Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
                  McCann,R., Waterston,R. and Wilson,R.
                  Public Soybean EST Project
                  Unpublished (1999)
                  Contact: Shoemaker R/Public Soybean EST Project
                  Public Soybean EST Project
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: est@watson.wustl.edu
                  When it has been determined, an EST from the other end of this
                  clone is listed in the 'Other ESTs on clone' field. This clone is
                  available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
                  57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
                  Seq primer: -40RP from Gibco
                  High quality sequence stop: 421.
                  Location/Qualifiers
FEATURES           source
                   1..501
                   /organism="Glycine max"
                   /mol_type="mRNA"
                   /cultivar="Raiden"
                   /db_xref="taxon:3847"
                   /clone="SOYBEAN CLONE ID: Gm-cl061-3780"
                   /tissue_type="mature flowers of field grown plants"
                   /lab_host="DH10B"
                   /clone_lib="Gm-cl061"
                   /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
                   XhoI; The cDNA library was constructed from mRNA isolated
                   from mature flowers of field grown plants for the cultivar
                   Raiden. Complementary DNA was synthesized from mRNA using
                   a primer consisting of a poly(dT) sequence with a XhoI
                   restriction site. EcoRI adapters were ligated to the
                   blunt-ended cDNA fragments followed by XhoI digestion. The
                   cDNA fragments were directionally cloned into the
                   EcoRI-XhoI restriction site of the pBluescript vector. The
                   ligated cDNA fragments were transformed into DH10B host
                   cells (GibcoBRL). This library was constructed in the
                   laboratory of Dr. Randy Shoemaker."
ORIGIN
Query Match      53.6%; Score 462.6; DB 4; Length 501;
Best Local Similarity 95.2%; Pred No. 1e-93;
Matches 477; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 20 CACAAGTCAATGGATAAAAAACAACAGTGTGAAGCGTCTCAAGATCTTCAAGTGGAGAAA 79
Db 1 CACAAAAAATCGATAAAAAACAACGTGTGCAACACGTCTCAAGATCTCTGAAGTGGAGAAA 60
QY 80 GGGCTTGGACATGGAAGAGACTTGATCTTGATGAACATATATGCAATCATGCGGAA 139
Db 61 GGACCTTGGACGATGGAAGAGACTTGATCTTGATGAACATATATGCAATCATGCGGAA 120
QY 140 GGTGTTTGGAACTCTTTGGCCAAAGCTGCTGTCTCTCAAAACGTAACGGAAAGAGTTGCCGG 199

```

```

Db 121 GGTGTTTGGAACTCTTTGGCCAAAGCTGCTGTCTCAAAACGTAACGGAAAGAGTTGCCGG 180
QY 200 CTAAGGTGGGTAAATTAACCTCCGTCTGATGTTTGAAGAGAGGAATATATACACCGAGAA 259
Db 181 CTAAGGTGGGTAAACTACCTCCGTCTGATGTTTGAAGAGAGGAATATATACACCGAGAA 240
QY 260 CAACCTTTTGATTATGGAGCTCCACGAAAGTGGGGAACAGGTGTCCTCAAAATTTGCCAAG 319
Db 241 CAACCTTTTGATCATGGAGCTTCACGAAAGTGGGGAACAGGTGTCCTCAAAATTTGCCAAG 300
QY 320 CATCTACCTGGAAGGACTGTAATAGATCAAGAACTATTTGGAGGACAAGGATCCAGAAG 379
Db 301 CATCTACTGCTAGACAGATAATAGATCAAGAACTATTTGGAGGACCAGGATCCAGAAG 360
QY 380 CACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTATAATATCTCGAGATAAATGAT 439
Db 361 CACATCAAGCAAGCTGAGAACTTTTCAGCAACAAATATAGTATAAATCTCTGAGATAAATGAT 420
QY 440 CACCAAGCTAGCACTAGCCATGTTTCCACCATGCTGAGCCCATGGAGATGATTTCTCCA 499
Db 421 CACCAAGCTAGCACTAGCCATGTTTCTACCATGCTGAACCCATGGAGACCTATTTCTCCA 480
QY 500 CCCTGTTTATCAAGGAATGTTA 520
Db 481 CCCTTTTATCAAGGAATGTTA 501

RESULT 5
BE058947          450 bp      mRNA      linear      EST 14-JUL-2004
LOCUS             sm23b01.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION        Gm-cl016-12290 5' similar to TR:Q927D5 Q927D5 PUTATIVE
                  TRANSCRIPTION FACTOR ;, mRNA sequence.
ACCESSION         BE058947
VERSION           BE058947.1 GI:8403313
KEYWORDS          EST.
SOURCE            Glycine max (soybean)
ORGANISM          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                  Glycine.
REFERENCE         1 (bases 1 to 450)
AUTHORS           Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
                  Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
                  Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
                  Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
                  Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
                  McCann,R., Waterston,R. and Wilson,R.
                  Public Soybean EST Project
                  Unpublished (1999)
                  Contact: Shoemaker R/Public Soybean EST Project
                  Public Soybean EST Project
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: est@watson.wustl.edu
                  When it has been determined, an EST from the other end of this
                  clone is listed in the 'Other ESTs on clone' field. This clone is
                  available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
                  57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
                  Seq primer: -40RP from Gibco
                  High quality sequence stop: 411.
                  Location/Qualifiers
FEATURES           source
                   1..450
                   /organism="Glycine max"
                   /mol_type="mRNA"
                   /cultivar="Williams 82"
                   /db_xref="taxon:3847"
                   /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-12290"
                   /tissue_type="immature flowers of field grown plants"

```

/lab_host="Xl10-Gold"
 /clone_lib="Gm-cl016"
 /note="vector: pBluescript II XR; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from immature flowers of field grown plants. The cDNA
 library was prepared using the Stratagene pBluescript II
 XR library construction kit. Complementary DNA was
 synthesized from mRNA using a primer consisting of a poly
 (dT) sequence with a XhoI restriction site. EcoRI adapters
 were ligated to the blunt-ended cDNA fragments followed by
 XhoI digestion. The cDNA fragments were directionally
 cloned into the EcoRI-XhoI restriction site of the
 pBluescript vector. The ligated cDNA fragments were
 transformed into Xl10-Gold host cells. This library was
 constructed by Dr. Randy Shoemaker and Dr. John
 Erpelnding."

ORIGIN

Query Match 51.5%; Score 444.4; DB 2; Length 450;
 Best Local Similarity 98.9%; Pred. No. 1.3e-89;
 Matches 445; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 304 GTCCAAATTCGCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAGAACTATTGGAG 363
 Db 1 GTCCAAATTCGCAAGCATCTACCTGGAAGGACTTATAAGATCAAGAACTATTGGAG 60

QY 364 GACAAGGATCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAA 423
 Db 61 GACAAGGATCAGAAGCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAA 120

QY 424 TTCTGAGATAAATGATCACAAGCTAGACATGAGCCATGTTTCCACATGAGTGGCCCAT 483
 Db 121 TTCTGAGATAAATGATCACAAGCTAGACATGAGCCATGTTTCCACATGAGTGGCCCAT 180

QY 484 GGAGATCTTCTCCACCTGTTATCAAGGAATCTTAGGCCATTTTCACTCAGTTCCC 543
 Db 181 GGAGATCTTCTCCACCTGTTATCAAGGAATCTTAGGCCATTTTCACTCAGTTCCC 240

QY 544 TACAATTAATCTGATCAATCACTGTTGTGTACCAATGACAAACAACTAACTATTG 603
 Db 241 TACAATTAATCTGATCAATCACTGTTGTGTACCAATGACAAACAACTAACTATTG 300

QY 604 GAGCATGGAGATAGCTGGTCAATGCAATTAATCAATGAGCGGTGATTAATATATCAAGATA 663
 Db 301 GAGCATGGAGATAGCTGGTCAATGCAATTAATCAATGAGCGGTGATTAATATATCAAGATA 360

QY 664 AATCTAAGTTTCAAGTTCATAGGCTGGAATGCTCTTGGATTAACATATTATGG 723
 Db 361 AATCTAAGTTTCAAGTTCATAGGCTGGAATGCTCTTGGATTAACATATTATGG 420

QY 724 GTTTGTTTATATAAGTAGTTGGATGTTGG 753
 Db 421 GTTTGTTTATATAAGTAGTTGGATGTTGG 450

RESULT 6

AI930997
 LOCUS
 DEFINITION B445h07.v1 Gm-cl015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl015-278 5' similar to TR:Q39028 Q39028 ATMYB2.1; mRNA
 sequence.

ACCESSION AI930997.1 GI:5666961

VERSION AI930997

KEYWORDS EST

SOURCE Glycine max (soybean)

ORGANISM

Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 523)

REFERENCE

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,

Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.

TITLE

JOURNAL

COMMENT

Public Soybean EST Project

Unpublished (1999)

Other ESTs: BE58316 corresponding to Gm-r1070-1759 (3')

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 422.

Location/Qualifiers

1..523

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Williams 82"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-278"

/tissue_type="Mature flowers, field grown plants"

/lab_host="Xl10-Gold"

/clone_lib="Gm-cl015"

/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:

XhoI; This cDNA library was constructed from mRNA isolated

from mature flowers of field grown plants. The cDNA

library was prepared using the Stratagene pBluescript II

XR cDNA library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were

transformed into Xl10-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Erpelnding."

ORIGIN

Query Match 51.3%; Score 442.8; DB 1; Length 523;
 Best Local Similarity 94.3%; Pred. No. 3e-89;
 Matches 459; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 8 CTCTATCACACACAAAGTCAATGGATATAAAAAACACAGGTGAAGAGCTCTCAAGATCCT 67
 Db 37 CTCTATCACACACAAAGTCAATGGATATAAAAAACAGTGTGCAACAGCTCTCAAGATCCT 96

QY 68 GAAGTGAGAAAAGGGCTTGGCAATATGGAGAGAGACTTGTATCTTTGATGAATATATGCA 127
 Db 97 GAAGTGAGAAAAGGGCTTGGAGAGAGAGACTTGTATCTTTGATGAATATATGCA 156

QY 128 AATCATGGGGAAGGTGTTTGGAACTCTTGGCCAAAGAGCTGCTGTCTCAACAGTAACGGA 187
 Db 157 AATCATGGGGAAGGTGTTTGGAACTCTTGGCCAAAGAGCTGCTGTCTCAACAGTAACGGA 216

QY 188 AAGAGTTGCCGGCTAAGGTGGCTAAATTAATCTCCGTCTCTGATGTTAGAGAGGGAATATT 247
 Db 217 AAGAGTTGCCGGCTAAGGTGGCTAAATTAATCTCCGTCTCTGATGTTAGAGAGGGAATATT 276

QY 248 ACACCCGAGGAAACAATTTTGTATATGAGAGCTCCACCAAGTGGGGGAAACAGGTGGTCC 307
 Db 277 ACACCCGAGGAAACAATTTTGTATATGAGAGCTTTCACCAAGTGGGGGAAACAGGTGGTCC 336

QY 308 AAAATTTGCCAAGCATCTACCTGGAGGACTGATTAATGAGATCAAGAACTATTGGAGACA 367
 Db 337 AAAATTTGCCAAGCATCTACCTGGAGGACTGATTAATGAGATCAAGAACTATTGGAGACC 396

QY 368 AGGATCCAGAACCATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCT 427
 |||||
 Db 397 AGGATCCAGAACCATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCT 456

QY 428 GAGATAAATGATACCAAGCTAGCACTAGCCATGTTTCCACATGGCTGAGCCCATGAG 487
 |||||
 Db 457 GAGATAAATGATACCAAGCTAGCACTAGCCATGTTTCCACATGGCTGAGCCCATGAG 516

QY 488 ATGTATT 494
 |||||
 Db 517 ACCTATT 523

RESULT 7
 AW459279
 LOCUS
 DEFINITION shz2h08.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl016-5560 5', similar to TR:Q9ZTDS 09ZTDS 09ZTDS PUTATIVE TRANSCRIPTION
 FACTOR ;, mRNA sequence.

ACCESSION AW459279.1 GI:7029496

VERSION EST.

KEYWORDS Glycine max (soybean)

SOURCE Glycine max

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 431)

REFERENCE
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE
 JOURNAL
 COMMENT Other ESTs: BE820766 corresponding to Gm-r1070-4966 (3')

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. This clone is
 available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Insert Length: 723 Std Error: 0.00

Seq primer: -40RP from Gibco

High quality sequence scop: 411.

Location/Qualifiers

1..431

/organism="Glycine max"

/mol type="mRNA"

/culivar="Williams 82"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-5560"

/tissue type="Immature flowers of field grown plants"

/lab host="XL10-Gold"

/clone lib="Gm-cl016"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from immature flowers of field grown plants. The cDNA
 library was prepared using the Stratagene pBluescript II
 XR library construction kit. Complementary DNA was
 synthesized from mRNA using a primer consisting of a poly
 (dT) sequence with a XhoI restriction site. EcoRI adapters
 were ligated to the blunt-ended cDNA fragments followed by
 XhoI digestion. The cDNA fragments were directionally
 cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were
 transformed into XL10-Gold host cells. This library was
 constructed by Dr. Randy Shoemaker and Dr. John
 Erpelding."

ORIGIN

Query Match 49.3%; Score 425.4; DB 2; Length 431;
 Best Local Similarity 98.8%; Pred. No. 2.5e-85;
 Matches 426; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 319 GCATCTACCTGGAAGGACTGATATAGATCAAGAACTATTGGAGGCAAGATCCAGAA 378
 |||||
 Db 1 GCACCAGCTGGAAGGACTGATAATAGATCAAGAACTATTGGAGGCAAGATCCAGAA 60

QY 379 GCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCTGAGATAAATGA 438
 |||||
 Db 61 GCACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCTGAGATAAATGA 120

QY 439 TCACCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTTATCTCC 498
 |||||
 Db 121 TCACCAAGCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTTATCTCC 180

QY 499 ACCCTGTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTCAATAATTAATCTCTGA 558
 |||||
 Db 181 ACCCTGTATCAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTCAATAATTAATCTCTGA 240

QY 559 TCAATCCAGTGTGTACCAATGACAAACAAACATTAACCTATTGGAGCATGGAGGATAG 618
 |||||
 Db 241 TCAATCCAGTGTGTACCAATGACAAACAAACATTAACCTATTGGAGCATGGAGGATAG 300

QY 619 CTGTCGAATCAATTAATCTGAAACCGTGATTAATAATTAATCAAGATAAAACCTAAGTTTGA 678
 |||||
 Db 301 CTGTCGAATCAATTAATCTGAAACCGTGATTAATAATTAATCAAGATAAAACCTAAGTTTGA 360

QY 679 AGTTCCATAAGCTCGAATGCTCTGGATTAACATATATTGGTTGTTTATATAAG 738
 |||||
 Db 361 AGTTCCATAAGCTCGAATGCTCTGGATTAACATATATTGGTTGTTTATATAAG 420

QY 739 TAGTTGGATGT 749

Db 421 TAGTTGGATGT 431

RESULT 8

BM527774

LOCUS

DEFINITION

sal165903.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl061-4134 5' similar to TR:O49020 O49020 MYB-LIKE DNA-BINDING
 DOMAIN PROTEIN. ;, mRNA sequence.

ACCESSION BM527774

VERSION BM527774.1 GI:18733722

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 552)

REFERENCE

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 426.
Location/Qualifiers
1. 552

FEATURES

source
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Raiden"
/db_xref="taxon:3847"
/clones="SOYBEAN CLONE ID: Gm-cl061-4134"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."
ORIGIN

Query Match 49.2%; Score 424.4; DB 4; Length 552;
Best Local Similarity 91.7%; Pred. No. 4.2e-85;
Matches 506; Conservative 2; Mismatches 28; Indels 16; Gaps 5;
Qy 173 CTCAACGTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTCTGATGTT 232
Db 1 CTCAACGTAACGGAAGAGTTCGGCTAAGTGGCTAAATTAACCTCGCTCTGATGTT 60
Qy 233 AGAAGAGGGAATATTACACCGGAGAAACACTTTTGTATTGAGCTCCAGCAAGTGG 292
Db 61 AGAAGAGGGAATATTACACCGGAGAAACACTTTTGTATTGAGCTCCAGCAAGTGG 120
Qy 293 GGAACAGGTGGTCCAAAATTGCCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAG 352
Db 121 GGAACAGGTGGTCCAAAATTGCCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAG 180
Qy 353 AACTATTGGAGGACCAAGATCCAGACATCAAGCAAGCTGAGACTTTCAGCAACAG 412
Db 181 AACTATTGGAGGACCAAGATCCAGACATCAAGCAAGCTGAGACTTTCAGCAACAA 240
Qy 413 AGTAGTAATATCTGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATG 472
Db 241 ATTAGTAATATCTGAGATAAATGATCACCAGCTAGCACTAGCCATGTTTCCACCATG 300
Qy 473 GCTGAGCCCATGGAGATGTTATCTCCACCTGTTTATCAAGGAATGTTAGAGCCATT---T 529
Db 301 GCTGAGCCCATGGAGACCTATTCTCCACCTTTTATCAAGGAATGTTAGAGCCATTTCT 360
Qy 530 TCAACTAGTTCCTTACATAATTAATCCTGATCAATCCAGTTGTTTGTACATGACAAAC 589
Db 361 TCAATTGAGTTCCTTACATAATTAATCCTGATCAATCCAGTTGTTTGTACATGACAAAC 420
Qy 590 AA---CATTAATCTATGGAGATGAGGATAGCTAGTGGTCAATGCAATTAAGCGGTGAT 646
Db 421 AACAGCATTAATCTATGGAGATGAGGATATCTGGTCAATGCAATTAAGCGGTGAT 480
Qy 647 TAAATAT-----TATCAAGATAAAACCTTAAGTTTG---AAGTTTCCAT--AAGGCTGGAA 696
Db 481 TAAATATGATATATCAAGATAAAACCTTAATTTCTTGTTATTAAGTTTCCATAAAACACTGGAA 540
Qy 697 TGCTTTGGATT 708
|||||

Db 541 TGTCCTGTGGCTT 552

RESULT 9

BM527664

LOCUS

DEFINITION

BM527664

ACCESSION

BM527664

VERSION

BM527664.1

KEYWORDS

EST

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

REFERENCE

AUTHORS

1 (bases 1 to 568)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

TITLE

JOURNAL

COMMENT

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 422.

FEATURES

source

1. 568

Location/Qualifiers

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Raiden"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl061-4003"

/tissue_type="mature flowers of field grown plants"

/lab_host="DH10B"

/clone_lib="Gm-cl061"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 46.5%; Score 401.2; DB 4; Length 568;
Best Local Similarity 85.6%; Pred. No. 7.2e-80;
Matches 471; Conservative 0; Mismatches 73; Indels 6; Gaps 2;
Qy 13 TCACACACACAGTCAATGGATGATAAAACACAGTGTAAAGACGCTCAAGATCTCTGAAGT 72
Db 13 TCCTCACACAAATAATATGACAAGA---AGCTGGCAACACGCTCTCATGATCTCTGAAGT 69
Qy 73 GAGAAAGGGCCCTTGGCAATGGAAGACTTGATCTTGTAGTGAATATATTGCAAAATCA 132

Db 70 GAGAAAGGGCCATGGCAATGGAAGAGACTTAATCTTGATCACTATATTGGCAATCA 129
Qy 133 TGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGTCTCAAACTGAACGGAAGAG 192
Db 130 CGGGGAAGGGGTTTGGAACTCTTTGGCCAAAGCTGCTGTCTCAAACTGAACGGAAGAG 189
Qy 193 TTGCGGGCTAAAGGTGGCTAAATTAACCTCCGCTCCTGATGTTAGAAGAGGAATATTACACC 252
Db 190 TTGCGGGCTCCGGTGGCTAACTACCTCCGCTCCTGATGTTAGAAGAGGGAATATTACACC 249
Qy 253 CGAGAAACAATTTTGATATTGAGACTCCAGCAAAAGTGGGAAACAGGTGGTCCAAAAT 312
Db 250 CGAGAAACAGCTTTTGATCATGGAATTTATGCAAAAGTGGGAAACAGGTGGTCCAAAAT 309
Qy 313 TGCCAAGCATCTACCTCGAAGGACTGATAATGAGATCAAGAACTATTTCGGAGGCAAGGAT 372
Db 310 TGCCAAGCATCTACCCGGAAGACTGATAATGAGATTAAGAACTACTCGAGGCAAGGAT 369
Qy 373 CCAGAAGCACATCAAGCAAGCT--GAGAACTTTTCAGCAACAGAGTAGTAATAATTCTGA 429
Db 370 CCAGAAGCACCTCAAGCAAGCTTCAGCAGCTTCAGCAACAGAGTAGTAATTCTGAGAT 429
Qy 430 GATAAATGATCACCAAGCTAGACATAGCCATGTTTCCACCATGGCTGAGCCCATGGAGAT 489
Db 430 AATTTATCATCCCAAGCTTGCACTAGCAAGTGTCCACCATGGCGCAGCCCATAGAAAC 489
Qy 490 GTATTCTCCACCTGTTATCAAGGAATCTAGAGCCATTTCAACTCAGTTCCCTACAAAT 549
Db 490 GTATTCTCCACCCAGTTATCAAGGAATGTTAGATCCATTTTCAATTTCAGTTCCCAACAA 549
Qy 550 TAATCCTGAT 559
Db 550 TCCATCAT 559

RESULT 10
BM527508
LOCUS
DEFINITION
sa162f08.yl Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1061-3735 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA
sequence.

ACCESSION
BM527508
VERSION
BM527508.1 GI:18733261

KEYWORDS
EST.

SOURCE
Glycine max (soybean)

ORGANISM
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 544)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco

High quality sequence stop: 422.

FEATURES
source

Location/Qualifiers
1..544
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Raiden"
/db_xref="taxon:3847"
/clones="SOYBEAN CLONE ID: Gm-c1061-3735"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 46.4%; Score 400.6; DB 4; Length 544;
Best Local Similarity 86.3%; Pred. No. 9.7e-80;
Matches 455; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy 36 AAAAACAACAGTGTAAAGCGTCTCAAGTCTGAGTGAAGAAAGGGCTTTGGACATGG 95

Db 18 ACAAGAAGCTTGGCAACACCGTCTCATGATCTGAAGTGAGAAAGGGCCATGGCAATGG 77

Qy 96 AAGAAGACTTGATCTTGATGAATATATTGCAAAATCATGGGAAAGGTGTTTGGAACTCTT 155

Db 78 AAGAAGACTTAACTTGTGATCACCTATATTGCCATCACGGGAAGGGGTTTGGAACTCTT 137

Qy 156 TGGCCAAAGCTGCTGGTCTCAACGTAAACGTAACGAAAGAGTTGCCGGCTAGGTGGCTAAAT 215

Db 138 TGGCCAAAGCTGCTGGACTTAAAGCTACCGGAAAGAGTTGCCGGCTCCGGTGGCTAAAT 197

Qy 216 ACCTCCGCTCGATGTTAGAGAGGGATATTACACCGGAGGAACACTTTTGGATTATGG 275

Db 198 ACCTCCGCTCGATGTTAGAGAGGGATATTACACCGGAGGAACACTTTTGGATTATGG 257

Qy 276 AGCTCCACGCAAAAGTGGGAAACAGGTGCTCCAAAATTTGCCAAGCATCTACCTGGAAGGA 335

Db 258 AACTTCATGCAAAAGTGGGAAACAGGTGCTCCAAAATTTGCCAAGCATCTACCTGGAAGGA 317

Qy 336 CTGATAATGAGATCAAGAACTATTGGAGGAGCAAGATCCAGAACATCAAGCAAGCT- 394

Db 318 CTGATAATGAGATTAAGAACTACTCGAGGACCAAGGATCCAGAACACCTCAAGCAAGCTT 377

Qy 395 --GAGAACTTTCAGCAACAGAGTAGTAATAATTCTGAGATAAATGATCACCAGCTAGCA 452

Db 378 CCAGCAGCTTCCAGCAACAGAGTAGTAATAATTCTGAGATAAATGATCACCAGCTAGCA 437

Qy 453 CTAGCCATGTTTCCACCATGCTGAGCCCATGGAGATGATTTCTCCACCTGTTTATCAAG 512

Db 438 CTAGCCAGTGTCCACCATGCGCAGCCCATAGAAACCTATTCTCCACCCAGTTATCAAG 497

Qy 513 GAATGTTAGAGCCATTTTCAACTCAGTCTCCCTCAACAATTAATCCTGAT 559

Db 498 GAATGTTAGATCCATTTTCAATTTCAGTTTCCCAACAATAATCCTCATCAT 544

RESULT 11

BM527508

LOCUS

DEFINITION

ss45f11.yl Gm-c1061 Glycine max cDNA clone

Gm-c1061-1702 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA

sequence.

BM527508

ACCESSION

BM527508.1

GI:10235902

409 bp mRNA linear EST 13-JUL-2004
ss45f11.yl Gm-c1061 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1061-1702 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA

BE804790

DEFINITION

ss45f11.yl Gm-c1061 Glycine max cDNA clone

Gm-c1061-1702 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA

sequence.

BE804790

ACCESSION

BE804790.1

GI:10235902

```

KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 409)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
        Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
        Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
        Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
        Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
        McCann,R., Waterston,R. and Wilson,R.
        Public Soybean EST Project
        Unpublished (1999)
        Contact: Shoemaker R/Public Soybean EST Project
        Public Soybean EST Project
        Washington University School of Medicine
        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
        Tel: 314 286 1800
        Fax: 314 286 1810
        Email: est@watson.wustl.edu
        When it has been determined, an EST from the other end of this
        clone is listed in the 'Other ESTs on clone' field. This clone is
        available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
        57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
        Insert Length: 1068 Std Error: 0.00.

FEATURES
    source
        1..409
            /organism="Glycine max"
            /mol_type="mRNA"
            /cultivar="Raiden"
            /db_xref="taxon:3847"
            /clone="GENOME SYSTEMS CLONE ID: Gm-cl061-1702"
            /tissue_type="mature flowers of field grown plants"
            /lab_host="DH10B"
            /clone_lib="Gm-cl061"
            /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
            XhoI; The cDNA library was constructed from mRNA isolated
            from mature flowers of field grown plants for the cultivar
            Raiden. Complementary DNA was synthesized from mRNA using
            a primer consisting of a poly(dT) sequence with a XhoI
            restriction site. EcoRI adapters were ligated to the
            blunt-ended cDNA fragments followed by XhoI digestion. The
            cDNA fragments were directionally cloned into the
            EcoRI-XhoI restriction site of the pBluescript vector. The
            ligated cDNA fragments were transformed into DH10B host
            cells (GibcoBRL). This library was constructed in the
            laboratory of Dr. Randy Shoemaker."

ORIGIN
    Query Match 42.9%; Score 369.8; DB 2; Length 409;
    Best Local Similarity 95.7%; Pred. No. 8.4e-73;
    Matches 380; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 8 CTCTATCACACACAAAGTCAATGGATAAAAAACACAGCTGTAAAGCTCTCAAGATCCT 67
Db 13 CTCTATCACACACAAACAAATGGATAAAAAACACTGTGCACACAGTCTCAAGATCCT 72
Qy 68 GAAGTGAGAAAGGGCCCTTGGACAAATGGAAGAGCTTGATCTTTGATGAATATTGCA 127
Db 73 GAAGTGAGAAAGGACCTTGGACGATGGAAGAGCTTGATCTTTGATGAATATTGCA 132
Qy 128 AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGGA 187
Db 133 AATCATGGGAGAGTGTGGAACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGGA 192
Qy 188 AAGAGTTGCCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 247
Db 193 AAGAGTTGCCCGCTAAGGTGGCTAAATTAACCTCCGTCCTGATGTTAGAGAGGGAATATT 252
Qy 248 ACACCCGAGGGAACAACATTTTGATTTATGGAGCTCCACGCAAGAGTGGGGAACAGGTGTCTC 307

```

```

Db 253 ACACCCGAGGGAACAACATTTTGATCTGGAGCTTCACGCAAAAGTGGGAAAACAGGTGTCTC 312
Qy 308 AAAATTGCCAAGCATCTTACCTGGAAGGACTGATAATGAGATCAAGAACTATTGGAGACA 367
Db 313 AAAATTGCCAAGCATCTTACCTGTAAGACAGATTAAGATCAAGATCTATTGGAGGACC 372
Qy 368 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTC 404
Db 373 AGGATCCAGAAGCACATCAAGCAAGCTGAGAACTTTC 409

RESULT 12
LOCUS BM732539
DEFINITION BM732539.1 GI:19053872
            Glycine max (soybean)
            Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
            1 (bases 1 to 560)
            Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
            Public Soybean EST Project
            Unpublished (1999)
            Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            When it has been determined, an EST from the other end of this
            clone is listed in the 'Other ESTs on clone' field. This clone is
            available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
            57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 430.

FEATURES
    Location/Qualifiers
        1..560
            /organism="Glycine max"
            /mol_type="mRNA"
            /cultivar="Raiden"
            /db_xref="taxon:3847"
            /clone="SOYBEAN CLONE ID: Gm-cl061-5223"
            /tissue_type="mature flowers of field grown plants"
            /lab_host="DH10B"
            /clone_lib="Gm-cl061"
            /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
            XhoI; The cDNA library was constructed from mRNA isolated
            from mature flowers of field grown plants for the cultivar
            Raiden. Complementary DNA was synthesized from mRNA using
            a primer consisting of a poly(dT) sequence with a XhoI
            restriction site. EcoRI adapters were ligated to the
            blunt-ended cDNA fragments followed by XhoI digestion. The
            cDNA fragments were directionally cloned into the
            EcoRI-XhoI restriction site of the pBluescript vector. The
            ligated cDNA fragments were transformed into DH10B host
            cells (GibcoBRL). This library was constructed in the
            laboratory of Dr. Randy Shoemaker."

ORIGIN

```

Query Match	42.8%;	Score 369.6;	DB 4;	Length 560;
Best Local Similarity	80.0%;	Pred. No. 9.5e-73;		
Matches 448;	Conservative	0;	Mismatches 109;	Indels 3; Gaps 1;
QY	54	CGTCTCAAGATCCTGAAGTGAAGAAAGGGCCCTTGACAAATGGAGAAAGACTTGATCTTGA	113	
DB	4	CATCTCATGATCCTGAAGTGAAGAAAGGACCATGGACCATGGAAGAGACTTGATCTTGA	63	
QY	114	TGAACATATTGCAAAATCATGGGAAGGTGTTTGGAACTCTTTGGCCAAAAGCTGCTGGTC	173	
DB	64	TAAACTATATTGCAAAATCAGGTGAAGGTGTTTGGAACTCTTTAGCCAAAAGCTTCTGGTC	123	
QY	174	TCAAACGTAACGGAAGAGTGTCCGGCTTAAGTGGCTGCTAAATTAACCTCCGTCCTGATGTTA	233	
DB	124	TCAAACGAACGGGAAGAGTGTGCACTCCGTTGGCTTAACTACTCTTCGTCCTGATGTTA	183	
QY	234	GAAGAGGGAATATTACACCCGAGGAACAACTTTTGAATATGAGTCTCCAGCAAAAGTGGG	293	
DB	184	GAAGAGGAATTTTACACCCGAGGAACAGCTTTTGTATCATAGAACTTTCATGCAAAAGTGGG	243	
QY	294	GAACAGGTGGTCCAAAATTGCCAAGCATCTACCTGGAAGGACTGATAATGAGATCAAGA	353	
DB	244	GCAATAGTGGTCCAAAATTGCAAGCATCTTCCAGGAAGAACTGCAATGAGATTAGA	303	
QY	354	ACTATTGGAGGACAAAGGATCCAGAAAGCACATCAAGCAAGCTTGAGAACTTTTCAGCAACAGA	413	
DB	304	ACTTCTGGAGAACAGGATCCAAAGCACATTAAGCAAGCTGAGACTTTCACAAACATG	363	
QY	414	GTAGTAATAATTCTGAGATAAATGATACCAAGCTAGCACTAGCATTGTTCCACCATGG	473	
DB	364	GTAATTCAGAGAATAATGATCATCAAGCAAGCACTAGTACTAGCAAAAGTGTCCACCATGG	423	
QY	474	CTGAGCCCATGAGATGATTTCTCCACCCTGTATCAAGGAATGTTAGAGCCATTTTCAA	533	
DB	424	CACATCCAAATGAGACTTTCCTCCACCCTCATACCAAGAACTTTTGAGCCATTCAC	483	
QY	534	CTCAGTTCCTTACAAATTAATCCTGATCAATCCAGTGTGTTGACCAATGACAAACAACA	593	
DB	484	CTCAATTCCTTACAATCA---CTGATCAATCAAGTTGTTGTACCACCACCAACGACAACA	540	
QY	594	TTAACTATTGGAGCATGGAG	613	
DB	541	ACAACATATTGGAGCATCGAG	560	
RESULT 13				
BQ146360				
LOCUS				
DEFINITION	BQ146360 635 bp mRNA linear EST 24-APR-2002			
ACCESSION	NF047E06FL1F1050 Developing flower Medicago truncatula cDNA clone			
VERSION	NF047E06FL 5', mRNA sequence.			
KEYWORDS	BQ146360.1 GI:20283419			
SOURCE	EST:			
ORGANISM	Medicago truncatula (barrel medic)			
REFERENCE	Medicago truncatula			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.			
TITLE	1 (bases 1 to 635)			
JOURNAL	Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.			
COMMENT	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula flower library Unpublished (2001)			
	Contact: May GD			
	Plant Biology Division			
	The Samuel Roberts Noble Foundation			
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA			
	Tel: 580 224 6650			
	Fax: 580 224 6692			
	Email: gdmay@noble.org			
	Insert Length: 635 Std Error: 0.00			

Plate: 047	row: E	column: 06		
Seq primer:	TCACACAGGAACACGCTATGAC.			
Location/Qualifiers				
1. .635				
/organism="Medicago truncatula"				
/mol_type="mrna"				
/db_xref="taxon:3880"				
/clone="NF047E06FL"				
/cissue_type="Developing flowers"				
/dev_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."				
/note="Developing flower"				
/note="Vector: Lambda Zap: cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using EXaseist helper phage and the E. coli strain XLI-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."				
ORIGIN				
Query Match	42.7%;	Score 368.6;	DB 5;	Length 635;
Best Local Similarity	78.0%;	Pred. No. 1.6e-72;		
Matches 486;	Conservative	0;	Mismatches 124;	Indels 13; Gaps 3;
QY	8	CTCTATTCACACACAAGTCAATGGATATAAAAAACAACAGTGTGAAGCGTCTCAAGATCCT	67	
DB	13	CTCTCATATACATATATAAAAAATAATGGATAAAAAACCATGCAACTCATCTCAAGATCCT	72	
QY	68	GAAGTGAGAAAAGGCCTTGGACAATCGNAGAGACTTGATCTTGATGAACTATATTGCA	127	
DB	73	GAAGTGAGAAAAGGCCATGGACCATGGAAGAAGACTTGAATTTTAATCAATTTATATTGCA	132	
QY	128	AATCATGCGGGAAGGTGTTTGGAACTCTTTTGGCCAAAGCTGCTGCTCTCAACCGTAAACGGA	187	
DB	133	AATCATGTTGAAGGTGTTTGGAAATTCCTTAGCCAAAGCTGCTGCTCTTAAACCGTACAGGA	192	
QY	188	AAGAGTTCGCGCTAAGGTGGCTAAATTAACCTCGTCTCATGTTAGAGAGGGAATATT	247	
DB	193	AAAAGTTGCAGGCTTCGATGTTTAAACTACCTTCGTCAGATGTTAGAGAGGGAATATT	252	
QY	248	ACACCCGAGGAACAACATTTTTCATTTATGGAGCTCCACGCAAGTGGGAAACAGGTGTGTC	307	
DB	253	ACACCTGAGGAACAACATTTTTCATGGAACCTTCATGGAAGTGGGGAATAGTGTGTC	312	
QY	308	AAAATTGCCAAGCATCTACCTGGAAGGACTGATTAATGAGATCAAGAACTATTGGAGGACA	367	
DB	313	AAAATTGCCAAGCATCTTCCAGGAAGAACTGACAATGAGATTAAGAATTTTGGAGGACT	372	
QY	368	AGGATCCAGAGACATCAACAGCAAGCTTTCAGCAACAGAGT-----AGTAAT	421	
DB	373	AGGATACAAAAGCACTTAAGCAAGTGTGATCCTCACCAAAATAATTTCCACAAATG	432	
QY	422	AAATTCGAGATAAATGATCACCAGCTAGCACTAG-----CATGTTTCCACCATGGCTGA	477	
DB	433	AGTTTAGAGATAAATGATCATCATCAATCAAGTAGTAGCCCAAGTTTCCCAACATTGTAGA	492	
QY	478	GCCCATGAGAGATGATTTCTCCACCTGTATCAAGGAATGTTAGAGCAATTTTCAACTCA	537	
DB	493	GCCAAATGGAATCTTACTCTCCACCTTCATACCAAGGAACCTTTGGAGGCATTTCCAGCTCA	552	
QY	538	GTTCCTCTACAAT---TAATCCTGATCAATCCAGTTGTTGTACCAATGACAAACAACAT	594	
DB	553	ATTTCCAAACATCAATGATCATCTCCCTCAAGTTGTTGTACAAATGACAAACAATACAA	612	
QY	595	TAACTATTGGAGCATGGAGGATA	617	
DB	613	CAATTTATTGGAGCATGGAGGATA	635	

Plate: 047 row: E column: 06
Seq primer: TCACACAGAAACAGCTATGAC.
Location/Qualifiers
1. .635
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF047E06FL"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."
/clone_lib="Developing flower"
/note="Vector: Lambda Zap; cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XLI-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 42.7%; Score 368.6; DB 5; Length 635;
Best Local Similarity 78.0%; Pred. No. 1.6e-72;
Matches 486; Conservative 0; Mismatches 124; Indels 13; Gaps 3;
Qy 8 CTCTATCACACACACAAGTCAATGGATAAAAAACAACAGTGTAGAGCTCTCAAGATCCT 67
Db 13 CTCTCATATACATATAAAAAAATAATGGATAAAAAAACCATGCAACTCATCTCAAGATCCT 72
Qy 68 GAAGTGAGAAAGGGCCCTTGACAAATGGAAGAACTTGATCTTGTAGAACTATATTGCA 127
Db 73 GAAGTGAGAAAGGGCCCATGGACCATGGAAGAACTTGATTTTAATCAATATTATTGCA 132
Qy 128 AATCATGGGGAAGGTGTTTGGAACTCTTTGGCCAAAAGCTGCTGGTCTCAAAACGTAACGGA 187
Db 133 AATCATGGTGAAGGTGTTTGGAACTCTTTAGCCAAAGCTGCTGCTTAAACGTACAGGA 192
Qy 188 AAGAGTTCGGGCTAAAGTGGCTAAATTAACCTCGTCTGTATGTAGAGAGGGAATATT 247
Db 193 AAAAGTTGCAGGCTTCGATGGTTAAACTACCTTCGTCAGATGTTAGAGAGGGAATATT 252
Qy 248 ACACCCGAGGAACAACCTTTTGAATATGGAGCTCCACGCAAGTGGGAAACAGAGTGCTCC 307
Db 253 ACACCTGAGGAACAACCTTTTGAATATGGAGCTTCATGCAAGTGGGGAATAGGTGGTCC 312
Qy 308 AAAATTGCAAGCATCTACCTGGAAGGACTGATATAGAGATCAAGAACTATTGGAGGACA 367
Db 313 AAAATTGCAAGCATCTTCAGGAAGAACTGACAAATGAGATTAAGATTTTGGAGGACT 372
Qy 368 AGGATCCAGAGGACATCAAGCAAGCTGAGAACTTTTCAGCAACACAGAGT-----AGTAAT 421
Db 373 AGGATACAAAAGCACATTAAGCAAGTGTATCTCTCACCAAAATAATTTTCCAAACAATG 432
Qy 422 AATTCTGAGATAAATGATCACCAGCTAGCACTAG-----CCATGTTTCCACCATGGCTGA 477
Db 433 AGTTTAGAGATAAATGATCATCATCAATCAAGTAGTAGCCCAAGTTTCCAAACATTGTAGA 492
Qy 478 GCCCATGAGATGTATTCTCCACCCTGTATTCAAGGAATGTTAGAGGCATTTCACACTCA 537
Db 493 GCCAATGGAATCTTACTCTCCACCTTCATACCAAGGAACCTTTGGAGCCATTTCAGACTCA 552
Qy 538 GTTCCCTCAAAAT---TAATCTGATCAATCCAGTTGTTGTACCAATGACACAAACAACAT 594
Db 553 ATTTCCAAACAATCAATGATCATCATCCCTCAAGTTGTTGTACAAATGACAAACAATAACA 612
Qy 595 TAACCTATTGGAGCATGGAGGATA 617
Db 613 CAATTATTGGAGCATGGAGGATA 635

```

RESULT 14
BE057370
LOCUS
DEFINITION
  sn02c04.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
  Gm-c1015-8383.5; similar to TR:O49020 O49020 MYB-LIKE DNA-BINDING
  DOMAIN PROTEIN. ; mRNA sequence.
ACCESSION
VERSION
  BE057370
KEYWORDS
SOURCE
  BE057370.1 GI:8401736
  EST.
  Glycine max (soybean)
ORGANISM
  Glycine max
REFERENCE
  1 (bases 1 to 476)
  Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
  Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
  Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
  Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
  Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
  McCann,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  When it has been determined, an EST from the other end of this
  clone is listed in the 'Other ESTs on clone' field. This clone is
  available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
  57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
  Insert Length: 909 Std Error: 0.00
  Seq primer: -40RP from Gibco
  High quality sequence stop: 420.
FEATURES
  source
    1..476
    /organism="Glycine max"
    /mol_type="mRNA"
    /cultivar="Williams 82"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-8383"
    /tissue_type="Mature flowers, field grown plants"
    /lab_host="XL10-Gold"
    /clone_lib="Gm-c1015"
    /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
    XhoI; This cDNA library was constructed from mRNA isolated
    from mature flowers of field grown plants. The cDNA
    library was prepared using the Stratagene pBluescript II
    XR cDNA library construction kit. Complementary DNA was
    synthesized from mRNA using a primer consisting of a poly
    (dT) sequence with a XhoI restriction site. EcoRI adapters
    were ligated to the blunt-ended cDNA fragments followed by
    XhoI digestion. The cDNA fragments were directionally
    cloned into the EcoRI-XhoI restriction site of the
    pBluescript vector. The ligated cDNA fragments were
    transformed into XL10-Gold host cells. This library was
    constructed by Dr. Randy Shoemaker and Dr. John
    Erpelding."
ORIGIN
  Query Match 42.1%; Score 363.6; DB 2; Length 476;
  Best Local Similarity 87.8%; Pred. No. 2.1e-71;
  Matches 409; Conservative 0; Mismatches 54; Indels 3; Gaps 1;

  48 GTAAGAGCTCTCAAGATCTCTGAAGTGAAGAGGGCCCTTGACAAATGGAAGACTTGA 107
  |||||
  Db GCAACAGCTCTCATGATCTCTGAAGTGAAGAGGGCCATGACAAATGGAAGACTTAA 70

```

```

Qy 108 TCTTGATGAACATATATTGCAAAATCATGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTG 167
  |||||
Db 71 TCTTGATCACCTATATTGGCCAAATCACGGGAAGGGTTTGGAACTCTTTGGCCAAAGCTG 130

Qy 168 CTGGTCTCAAAACGTAACGGAAGAGTTGCCGGCTAAAGTGGCTAAATACCTCCGTCCTG 227
  |||||
Db 131 CTGGACTTAAACGCTACCGGAAGAGTTGCCGGCTCCGGTGGCTAAACCTACCTCCGTCCTG 190

Qy 228 ATGTTAGAAGAGGGAATATTACACCCGAGGAACAACATTTTGGATTATGGAGCTCCACGCA 287
  |||||
Db 191 ATGTTAGAAGAGGGAATATTACACCCGAGGAACAACATTTTGGATTATGGAGCTCCACGCA 250

Qy 288 AGTGGGGAACAGAGTGGTCCAAAATTTGCCAAGCATCTTACCTGGAAGGACTGATTAATGAGA 347
  |||||
Db 251 AGTGGGGAACAGAGTGGTCCAAAATTTGCCAAGCATCTTACCTGGAAGGACTGATTAATGAGA 310

Qy 348 TCAAGAACTATTGGAGGACAAAGGATCCAGAGGCACATCAAGCAAGCT---GAGAACTTTC 404
  |||||
Db 311 TTAAGAACTACTGGAGGACAAAGGATCCAGAGGCACCTCAAGCAAGCTTCCAGCAAGCTTCC 370

Qy 405 AGCAACAGAGTAGTAATAATTCTGAGATAAATGATACCAAGCTAGCACTAGCCTAGCTTT 464
  |||||
Db 371 AGCAACAGAGTAGTAATAATTCTGAGATAAATGATACCAAGCTAGCACTAGCCTAGCTTT 430

Qy 465 CCACCATGGCTGAGCCCATCGAGATGTATTCTCCACCCTGTATTCA 510
  |||||
Db 431 CCACCATGGCTGAGCCCATCGAGATGTATTCTCCACCCTGTATTCA 476

```

```

RESULT 15
BM528383
LOCUS
DEFINITION
  eal57f09.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
  Gm-c1061-3330.5; similar to TR:Q39028 Q39028 ATMYB2. ; mRNA
  sequence.
ACCESSION
VERSION
  BM528383
KEYWORDS
SOURCE
  Glycine max (soybean)
ORGANISM
  Glycine max

```

```

REFERENCE
  1 (bases 1 to 542)
  Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
  Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
  Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
  Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
  Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
  McCann,R., Waterston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  When it has been determined, an EST from the other end of this
  clone is listed in the 'Other ESTs on clone' field. This clone is
  available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
  57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
  Seq primer: -40RP from Gibco
  High quality sequence stop: 423.
FEATURES
  source
    1..542
    /organism="Glycine max"
    /mol_type="mRNA"
    /cultivar="Raiden"
    /db_xref="taxon:3847"

```

```
/clone="SOYBEAN CLONE ID: Gm-cl061-3330"  
/tissue type="mature flowers of field grown plants"  
/lab host="DH10B"  
/clone lib="Gm-cl061"  
/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."
```

ORIGIN

Query Match	40.2%;	Score 347;	DB 4;	Length 542;
Best Local Similarity	78.3%;	Pred. No. 1.2e-67;		
Matches 416;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0;
Qy 28	AATGATATAAAACAAACAGTGTAGACGCTCTCAAGATCCTGAAGTGAGAAAAGGCCTTG	87		
Db 5	AATGGACAAAACCAACATCAACTCATCATCTCATGATCCTGAAGTGAGAAAAGGCACATG	64		
Qy 88	GACATGGAAGAAGACTTGTATCTTGATCAACTATATTCAAATCATGGGGAAGGTGTTG	147		
Db 65	GACCATGGAAGAGACTTGTATCTTGATCAACTATATTCGCAATCACGGTGAAGGTGTTG	124		
Qy 148	GAATCTTTTGGCCAAAGCTGTGCTCTCAACGCTAACCGGAAAGAGTTGCCCGCTAAGGTG	207		
Db 125	GAATCTCTTAGCCAAAGCTTCTGTGCTCTCAACGAAACGGGAAAGAGTTGTCGACTCCGTTG	184		
Qy 208	GCTAAATTACCTCCGTCCTGATGTTAGAAAGGGGAAATATTACCCGAGGGAACAACCTTTT	267		
Db 185	GCTAAACTACCTTCGTCCTGATGTTAGAAAGGGGAAACATTACCCGAGGGAACAACCTTTT	244		
Qy 268	GATTATGAGCTCCACGAAAGTGGGAAACAGGTGGTCCAAAATTGCCAAGCATCTACC	327		
Db 245	GATCATAGAACTTCATGCAAGTGGGCAATAGGTGGTCCAAAATTGCCAAGCATCTCC	304		
Qy 328	TGGAAGGACTGATAATGAGATCAAGAACTATTGGAGGACAAGGATCCAGAAGCACATCAA	387		
Db 305	AGGAAGAAGTGCACATGAGATTAGAACTTCTGGAGAACAAAGATCCAAAAGCACATTAA	364		
Qy 388	GCAAGCTGAGAACTTTCAGCAACAGAGTAGTAAATTTCTGAGATAAATGATCACCACAGC	447		
Db 365	GCAAGCTGAGACTTCACAACACATGGTAAATTCAGAGAATAATGATCATCAAGCAAGCAC	424		
Qy 448	TAGCACTAGCCATGTTTCCACCATGGCTGAGCCATGGAGATGATTCTCCACCCCTGTTA	507		
Db 425	TAGTACTAGCAAGTGTCCACCATGGCACATCCAAATGAGACTTCTCTCCACCCCTCAT	484		
Qy 508	TCAAGGAATGTTAGAGCCATTTTCAACTAGTTCCTTCAATTAATCCTGA	558		
Db 485	CCAAGGAACCTTTTGAGCCATTCACCTCAATTCCTCTACATCACTGATCA	535		

Search completed: September 28, 2005, 11:37:27
Job time : 3566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2005, 11:37:32 ; Search time 4143 Seconds
(without alignments)
2409.311 Million cell updates/sec

Title: US-10-659-869A-36

Perfect score: 1138
Sequence: 1 MDKQKQCTSQDPEVRKGPW.....NNINYSWEDSWSQLNGD 206

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10659869/runat_26092005_143445_11370/app.query.fasta_1.391
-DB=genEmbl -QWMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10659869@cgn 1.1 3731 @runat_26092005_143445_11370 -NCPU=3
-NO.MMAP -LARGQUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	909.5	79.9	1022	8 PSMYB26	Y11105 P sativum m
2	733.5	64.5	745	8 GHY54698	AY554698 Gerbera h
3	671	59.0	910	8 AY086615	AY086615 Arabidops
4	667	58.6	645	8 AY519632	AY519632 Arabidops

5	667	58.6	645	8	AF175987	Arabidops
6	666.5	58.6	681	8	AY519589	Arabidops
7	666.5	58.6	1027	8	AK118439	Arabidops
8	653.5	57.4	1033	6	AR439844	Sequence
9	653.5	57.4	1033	6	AB005888	Arabidops
10	592	52.0	621	8	AY519582	Arabidops
11	592	52.0	652	8	BT005574	Arabidops
12	592	52.0	894	8	AK118091	Arabidops
13	585	51.4	1127	8	AY088761	Arabidops
14	570	50.1	872	8	AB058642	Lilium hy
15	561.5	49.3	649	8	AF485900	Dendrobil
16	539	47.4	861	8	AY519568	Arabidops
17	537	47.2	903	6	AX653670	Sequence
18	532	46.7	990	6	AX654067	Sequence
19	532	46.7	990	6	AX699688	Sequence
20	532	46.7	990	6	AX699719	Sequence
21	532	46.7	1645	8	AK120551	Oryza sat
22	523.5	46.0	924	8	AY519635	Arabidops
23	520	45.7	945	6	AX654022	Sequence
24	520	45.7	1295	8	CPU33917	Craterostig
25	519.5	45.7	10543	8	ATAC009325	Arabidops
26	519	45.6	1421	8	AK103455	Oryza sat
27	518	45.5	858	6	AX653047	Sequence
28	518	45.5	1343	8	AY026332	Oryza sat
29	515.5	45.3	1117	8	AF510112	Craterost
30	515	45.3	1129	8	BT009018	Triticum
31	514.5	45.2	13242	8	AP003140	Oryza sat
32	514.5	45.2	145491	8	AP002883	Oryza sat
33	510	44.8	972	8	AY519584	Arabidops
34	510	44.8	1423	8	AF262733	Arabidops
35	509	44.7	1404	8	AK069082	Oryza sat
36	504.5	44.3	1077	6	AX653037	Sequence
37	500.5	44.0	1575	8	AK107424	Oryza sat
38	497	43.7	852	8	AY519558	Arabidops
39	497	43.7	1140	8	AF334815	Arabidops
40	485	42.6	732	8	AY519562	Arabidops
41	485	42.6	931	8	AY008377	Arabidops
42	483	42.4	1062	6	AX699667	Sequence
43	479	42.1	853	8	AF474132	Sorghum b
44	476.5	41.9	125993	8	AP003816	Oryza sat
45	476.5	41.9	130973	8	AP003813	Oryza sat

ALIGNMENTS

RESULT 1
PSMYB26
LOCUS PSMYB26 1022 bp mRNA linear PLN 09-FEB-1998
DEFINITION P.sativum mRNA for Myb-like protein (Myb26).
ACCESSION Y11105
VERSION Y11105.1 GI:1841474
KEYWORDS Myb-like protein; Myb26.
SOURCE Pisum sativum (pea)
ORGANISM Pisum sativum
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae;
Pisum.
REFERENCE 1
Uimari, A. and Strommer, J.
Myb26: a MYB-like protein of pea flowers with affinity for
promoters of phenylpropanoid genes
Plant J. 12 (6), 1273-1284 (1997)
MEDLINE 98112025
PUBMED 9450341
REFERENCE 2 (bases 1 to 1022)
Strommer, J.N.
Direct Submission
Submitted (05-FEB-1997) J.N. Strommer, University Of Guelph,
Molecular Biology And Genetics, Guelph, Ontario N1G 2W1, CANADA
JOURNAL
TITLE Location/Qualifiers
AUTHORS 1. .1022
JOURNAL /organism="Pisum sativum"
FEATURES
source

```
/mol_type="mRNA"  
/db_xref="taxon:3888"  
/tissue_type="flower bud"  
92...745  
/codon_start=1  
/product="Myb26"  
/protein_id="CAA71992.1"  
/db_xref="GI:1841475"  
/db_xref="GOA:P93474"  
/db_xref="UniProt/TREMBL:P93474"  
/translation="MDKKPCNSQDPEVRKGPWTMEEDLILINIANHGEVWNSLAK  
AAGKRTKSCRLRWLYLPDVRGNITPEEQLLIMELHSHKGNRWSKIACHLPGR  
DNEIKNFWRIQKHIOVDNPNQONFOKMSLEINDHHHPHOPSSOVSNLVEPM  
ETYSPTSQGTLEPPTQFPPTINDHQNCCANDNNNNWSMEDLWSQQLNGD"  
  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,63e-85 Length: 1022  
Score: 909.50 Matches: 174  
Percent Similarity: 83.18% Conservative: 9  
Best Local Similarity: 79.09% Mismatches: 20  
Query Match: 79.92% Indels: 17  
DB: 8 Gaps: 5  
  
US-10-659-869A-36 (1-206) x PSMYB26 (1-1022)  
  
Qy 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20  
Db ATGGACAAAAA---CCCTGCAACTCATCTCAAGATCTGGAAGTAGAGAAAGGCGCATGG 148  
  
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40  
Db ACCATGGAGAGAGACTTGATTTGATCAATATATATGCAATCATGTTGAAGGTGTTGG 208  
  
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60  
Db AATTCCTTAGCCAAAGCTGCTGCTTAAACGCACGGGAAAGAGTTGCAGGCTTCGATGG 268  
  
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluLeuLeu 80  
Db TTAACACTATCTTCGTCGCGATGTTAGAGAGAGAAATATTACACCTGAGGAACAACATTTTG 328  
  
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100  
Db ATCATGGAACTTCATTCAAAGTGGGAAATAGTGTGTCCTCAAAATTTGCAAGCATCTTCCA 388  
  
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120  
Db GGAAGAACTGACATGAGATTAAAGAAATTTTGGAGGACTAGGATACAAAAGCACATTAAG 448  
  
Qy 121 GlnAla-----GluAsnPheGlnGlnSerSerAsnAsnSerGluIle 135  
Db CAAGTTGATAACCCTAATCAACAAATTTTCAGCAAAATAATGAGT-----TTAGAGATA 502  
  
Qy 136 AsnAsp-----HisGlnAlaSerThrSerHisValSerThrMetAla 149  
Db AATGATCATCACCATCATCTCTCATCAACCAAGTAGTAGCAAGTTCTTAATTGGTA 562  
  
Qy 150 GluProMetGluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThr 169  
Db GAGCAATGGAACCTATTCTCCAACTTCATATCAAGGAACTTTGGAGCAATTTTCCAACT 622  
  
Qy 170 GlnPheProThrIleAsnProAsp-----GlnSerSerCysGlyThrAsnAspAsn 186  
Db CAATTTCCAACTATTAAATAATCATCATCAAACTCAAAATCAATTTGTCCTCAATGACAA 682  
  
Qy 187 AsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206  
Db AATAACAACAATATTGGAGCATGGAGGATATCTGTCATGCAATTAATCTCAATGAGAT 742  
  
RESULT 2  
GHY554698 745 bp mRNA linear PLN 07-JAN-2004  
LOCUS
```

Gerbera hybrid cv. 'Terra Regina' mRNA for MYB8 protein.
AJ554698 GI:40643881
myb8 gene, MYB8 protein.
Gerbera hybrid cv. 'Terra Regina'
Gerbera hybrid cv. 'Terra Regina'
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Mutisioideae;
Mutisieae; Gerbera.
1
Elomaa, P., Uimari, A., Mehto, M., Albert, V.A., Laitinen, R.A. and
Teeri, T.H.
Activation of anthocyanin biosynthesis in Gerbera hybrida
(Asteraceae) suggests conserved protein-protein and
protein-promoter interactions between the anciently diverged
monocots and eudicots
Plant Physiol. 133 (4), 1831-1842 (2003)
14605235
2 (bases 1 to 745)
Elomaa, P.
Direct Submission
Submitted (28-MAR-2003) Elomaa P., Department of Applied Biology,
University of Helsinki, P.O.Box 27, 00014 University of Helsinki,
FINLAND
Location/Qualifiers
1..745
/organism="Gerbera hybrid cv. 'Terra Regina'"
/mol_type="mRNA"
/cultivar="Terra Regina"
/db_xref="taxon:226891"
/tissue_type="petal"
1..745
/gene="myb8"
40..612
/gene="myb8"
/function="transcription factor"
/codon_start=1
/product="MYB8 protein"
/protein_id="CAD87008.1"
/db_xref="GI:40643882"
/translation="MEKRPNNSSQDPEVRKGPWTMEEDLILINIANHGEVWNSLAK
SAGKRTKSCRLRWLYLPDVRGNITPEEQLLIMELHSHKGNRWSKIACHLPGR
DNEIKNFWRIQKHIOVDNPNQONFOKMSQDQATMSSTQTCNVNPMETYSPPYDNF
QNFSPPPTETNENMWSMEDLWSQQLNSE"

ORIGIN
Alignment Scores:
Pred. No.: 8,15e-67 Length: 745
Score: 733.50 Matches: 146
Percent Similarity: 76.70% Conservative: 12
Best Local Similarity: 70.87% Mismatches: 29
Query Match: 64.46% Indels: 19
DB: 8 Gaps: 6

US-10-659-869A-36 (1-206) x GHY554698 (1-745)

Qy 3 LysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrpThrMet 22
Db GAAAAAGGCGCAACAATTTCTCGCAAGACGTCGAAAGTTAGAAAGGCGCATGGACGATG 102

Qy 23 GluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGlyValTrpAsnSer 42
Db GAAGNAGACTTGATTCCTTATAAATACATCGCAATCATCGGAAGGTGTTTGAACCTCT 162

Qy 43 LeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsn 62
Db CTTGCGATCCGAGGTCTTAAAAAGAACCCGAAAAAGTTGCCGCTTCGGTGGCTAAAC 222

Qy 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeuMet 82
Db TATCTCCGCGCAGATGTAAGGAGGGGTAAATTATACCCGAGAAACAACCTCTTGATCATG 282

Qy 83 GluLeuHisAlaLysTrpGlyAsnArgTrpSerLysLeuAlaLysHisLeuProGlyArg 102
 Db 283 GAGCTTCATGCGCAATGGGAAATAGATGGTCAAAAATAGCAAAATCTTCCAGGAGA 342
 Qy 103 ThrAspAsnGluLeuLysAsnTrpArgThrArgLysHisLeuGlnAla 122
 Db 343 ACAGATACGAGATTAAGAAATATTGGAGAACAGAAATTCAAAACATCAAGCAAGTT 402
 Qy 123 GluAsnPheGlnGlnSerSerAsnSerGluLeuAsnAspHisGlnAlaSerThr 142
 Db 403 GMAAATTTTGTAGTGGCAAACTCTTCAGGACAA-----AACGAC--CAAGCCACCATG 453
 Qy 143 SerHisValSerThr-----MetAlaGluProMetGluMetTyrSerProCysTyr 160
 Db 454 AGCTCCACACAAACATCAATGTTGTAATCCATGAAACGATATTCCTCCT--TAT 510
 Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrLeuAsnProAspGlnSerSer 180
 Db 511 GACAAAT---TTTCAGAAATTTTCAGTCCATTTCCAC----- 546
 Qy 181 CysCysThrAsnAspAsnAsnAsnLeuAsnTyrTrpSerMetGluAspSerTrpSerMet 200
 Db 547 -----GAAACAAATGAGACATGTTGGAGATGGAGGATTTATGTCCTCATG 591
 Qy 201 GlnLeuLeuAsnGlyAsp 206
 Db 592 CAGTTATTAAATAGCGAA 609

RESULT 3

AY086615 LOCUS AY086615 910 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 262460 mRNA, complete sequence.
 ACCESSION AY086615
 VERSION AY086615.1 GI:21405325

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 910)
 Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

AUTHORS Full-length messenger RNA sequences greatly improve genome
 annotation

TITLE Genome Biol. 3 (6), RESEARCH0029 (2002)

JOURNAL 22088475

MEDLINE 12093376

PUBMED 2 (bases 1 to 910)

REFERENCE Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.

AUTHORS Full-length cDNA from Arabidopsis thaliana

TITLE Unpublished

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 910)

AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available to TIGR and GenBank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the Ws or Laer ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Genset carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1..910

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="262460"

110..754

/codon_start=1

/product="putative transcription factor MYB24"

/protein_id="AAM63674.1"

/db_xref="GI:21554734"

/translation="MEKRSSSGSGSDAEVRKGPWTMEEDLILINIANHGBGVNWS
 LAKSAGLKRITGKSCRILNLYRPDVRNGNITPEEQITIMELHAKWNRVSKAKHLIP
 GRTDNEIKFWRTKIQYIIKSGETTTVGSQSEFINHHATTSHVNDTQETMDMYSP
 TTSYQHASNINQLNGYVPESSIMPLSVDSQSEYVWSVDDLWPMNIYNGN"

ORIGIN

Alignment Scores:

Pred. No.: 3,32e-60 Length: 910
 Score: 671.00 Matches: 133
 Percent Similarity: 72.90% Conservative: 23
 Best Local Similarity: 62.15% Mismatches: 50
 Query Match: 58.96% Indels: 8
 DB: 8 Gaps: 4

US-10-659-869A-36 (1-206) x AY086615 (1-910)

Qy 1 MetAspLysLysGlnGlnCys-----LysThrSerClnAspProGluValArgLysGly 18
 Db 110 ATGGAGAAAGAGAAAGTAGTGTGTTCTGGATCAGAGATGCGAGAGTGCAGAAAGGG 169
 Qy 19 ProTTrpMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
 Db 170 CCATGGACGATGGAGAGAGATTTTGATTCATCAATTATATCGCCATCATGTGAAGGT 229
 Qy 39 ValTrpAsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
 Db 230 GTTTGAAACTCTCGCCAAATCTGCAGGAGTAAACCGACCGCGGAAAGTTGCGGCTC 289
 Qy 59 ArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGln 78
 Db 290 CGGTGGCTGAACCTACCTCCACCTGATGTCGACGGGAAATATCACACCAAGAAACAG 349
 Qy 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHis 98
 Db 350 CTACCATCATGGAGACTTCATGCAAAATGGGAAATAGGTGGTCAAAAATTCGAAAGCAT 409
 Qy 99 LeuProGlyArgThrAspAsnGluLeuLysAsnTyrTrpArgThrArgLysGlnLysHis 118
 Db 410 TTACCAAGAGGAGCCGACATGAGATAAGAACTTTTGGAGAGCTAAGATCCAGAAATAC 469
 Qy 119 IleLysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluLeuAsnAspHis 138
 Db 470 ATCATCAAGAGCGGAGAAACGACGACGCGTTGGATCACAAAGCTCCGAGTTTATAACCAT 529
 Qy 139 GlnAlaSerThrSerHisVal---SerThrMetAlaGluProMetGluMetTyrSerPro 157
 Db 530 CATGCGACACAGGACCATGTCATGATGATGATCAAGAAACCATGGATATGTTCTCCA 589
 Qy 158 ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsn----- 175
 Db 590 ACGACGTCGTATCAACATGCCAGCAATATTAAATCAGCAGCTTAATTATGTAATTATGTG 649
 Qy 176 ProAspGlnSerSerCys-----CysThrAsnAspAsnAsnAsnIleAsnTyrTrp 192
 Db 650 CCTGAAATCCAGTTCGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 709
 Qy 193 SerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
 Db 710 AGCGTCGATGATCTTTGGCCCATGAATATATATATATATATATATATATATATATAT 751

```

RESULT 4
AY519632      AY519632      645 bp      mRNA      linear      PLN 07-FEB-2004
LOCUS      Arabidopsis thaliana MYB transcription factor (At5g40350) mRNA,
DEFINITION      complete cds.
ACCESSION      AY519632
VERSION      AY519632.1      GI:41619435
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE      1 (bases 1 to 645)
AUTHORS      Qu,L. and Gu,H.
TITLE      The MYB transcription Factor Family in Arabidopsis: A Genome-Wide
Cloning and Expression Pattern Analysis
JOURNAL
REFERENCE      2 (bases 1 to 645)
AUTHORS      Qu,L. and Gu,H.
TITLE      Direct Submission
JOURNAL      Submitted (07-JAN-2004) Life Sciences, National Laboratory of
Protein Engineering and Plant Genetic Engineering, Peking
University, Beijing 100871, China
FEATURES
source
1..645
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..645
/locus_tag="At5g40350"
1..645
/locus_tag="At5g40350"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AA510102.1"
/db_xref="GI:41619436"
/translation="MEKRESSGGSGDAEVRKGPWTWEEDLILINYIANHGEVWNS
LAKSAGLRETCKSLRWLNLVLRDPVREGNTPEOLTIMELHAKGNWSKIAHLPL
GRTDNEIKNFRWKIKQYIIKSGETTVGSSSFNFHATSHVNDTQETMDMYSF
TTSYOHASNINQNLNYGVPESGIMMPLSDVDSQYVSDVLDLWPMNLYNGN"
ORIGIN
Alignment Scores:
Pred. No.:      5,758-60      Length:      645
Score:      667.00      Matches:      132
Percent Similarity:      72.43%      Conservative:      23
Best Local Similarity:      61.68%      Mismatches:      51
Query Match:      58.61%      Indels:      8
DB:      8      Gaps:      4
US-10-659-869A-36 (1-206) x AY519632 (1-645)
Qy      1      MetAspLysLysGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18
Db      1      ATGGAGAAAAGAGAAAGTAGTGGTGGGCTCGGATCAGGAGATGCAGAGGTGAGAAAGGG 60
Qy      19      ProTTPThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
Db      61      CCATGGACGATGGAAGAAGATTGATTGCTTCATCAATTATATCGCCAATCATGGTGAAGGT 120
Qy      39      ValTTPAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db      121      GTTTGGAACTCTCTCGCCAAATCTCAGGACTAAACCGACCGGAAAAGTTGCCGGCTC 180
Qy      59      ArgTTPLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGln 78
Db      181      CGTGGCTGAACCTACTCCGACCTGATGTGCGACGGGAAATATCACACCAGAACAG 240
Qy      79      LeuLeuIleMetGluLeuHisAlaLysTTPGlyAsnArgTTPSerIleAlaLysHis 98
Db      241      CTCACCATCATGGAACCTTCATGCAAAATGGGGAAATAGTGTGTCACAAATTCGAAAGCAT 300

```

```

Qy      99      LeuProGlyArgThrAspAsnGluIleLysAsnTyrTTPArgThrArgIleGlnLysHis 118
Db      301      TTACAGGAAGAACCGACATAGATTAAGAAATTTTGGAGGACTAAGATCCAGAAATAC 360
Qy      119      IleLysGlnAlaGluAsnPheGlnGlnSerSerAsnSerGluIleAsnAspHis 138
Db      361      ATCATCAAGAGCGGAGAAACGACGACCGTTGGATCACAAGAGCTCCGAGTTTATAAACCAT 420
Qy      139      GlnAlaSerThrSerHisVal---SerThrMetAlaGluProMetGluMetTyrSerPro 157
Db      421      CATCGCAGACGAGCCATGTTCATGAATGATACTCAAGAAACCATGATGATGATTTCTCCA 480
Qy      158      ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsn----- 175
Db      481      ACGAGCTCGTATCAACATGCCAGCAATATTATCAGCAGCTTAATTATGGTAATTATGTG 540
Qy      176      ProAspGlnSerSerCys-----CysThrAsnAspAsnAsnAsnIleAsnTyrTrp 192
Db      541      CCTGAATCCGGTTCGATCATGCATTCATCTCTGTGATCAATCCGAACAAACCTATTGG 600
Qy      193      SerMetGluAspSerTTPSerMetGlnLeuLeuAsnGlyAsp 206
Db      601      ACGTCGATGATCTTTGGCCCATGAATATATATAATGGTAAT 642
RESULT 5
AF175987      645 bp      mRNA      linear      PLN 30-AUG-2001
LOCUS      Arabidopsis thaliana putative transcription factor (MYB24) mRNA,
DEFINITION      complete cds.
ACCESSION      AF175987
VERSION      AF175987.1      GI:5823306
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE      1 (bases 1 to 645)
AUTHORS      Kranz,H.D., Denekamp,M., Greco,R., Jin,H., Leyva,A., Meissner,R.C.,
Petroni,K., Urzaingui,A., Bevan,M., Martin,C., Smeekens,S.,
Tonelli,C., Paz-Ares,J. and Weisshaar,B.
TOWARDS FUNCTIONAL CHARACTERISATION OF THE MEMBERS OF THE R2R3-MYB
GENE FAMILY FROM ARABIDOPSIS THALIANA
Plant J. 16 (2), 263-276 (1998)
JOURNAL
MEDLINE      99056848
PUBMED      9839469
REFERENCE      2 (bases 1 to 645)
AUTHORS      Stracke,R., Werber,M. and Weisshaar,B.
TITLE      The R2R3-MYB gene family in Arabidopsis thaliana
JOURNAL      Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)
MEDLINE      21481677
PUBMED      11597504
REFERENCE      3 (bases 1 to 645)
AUTHORS      Stracke,R. and Weisshaar,B.
TITLE      Direct Submission
JOURNAL      Submitted (06-AUG-1999) Dept. Biochemie, Max-Planck-Institut fuer
Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
FEATURES
source
1..645
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Columbia"
1..645
/gene="MYB24"
1..645
/gene="MYB24"
/note="R2R3-MYB family member"
/codon_start=1
/product="putative transcription factor"
/protein_id="AA53092.1"
/db_xref="GI:5823307"
/translation="MEKRESSGGSGDAEVRKGPWTWEEDLILINYIANHGEVWNS

```

LAKSAGLRTGKSCRLRLWLYLRDPVRRGNITPEEQTLIMEHAKGNRRSKIAKHLPL
GRDNETKNFWRKIQYIIKSGSTTVGSGSEFINHATTSHVNDTQETWDMYSP
TTSYQHSANINQOLNYGVYPESGIMMPLSDQSEQNYVSDLLPMDNYNGN"

ORIGIN

Alignment Scores: 5.75e-60 Length: 645
Pred. No.: 667.00 Matches: 132
Score: 72.43% Conservative: 23
Percent Similarity: 61.68% Mismatches: 51
Best Local Similarity: 58.61% Indels: 8
Query Match: 8 Gaps: 4
DB:

US-10-659-869A-36 (1-206) x AF175987 (1-645)

Qy 1 MetAspLysLysGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18
Db 1 ATGGAGAAGAGAAAGTAGTGGTCTGGATCAGAGAGTGCAGAGGTGAGAAAGGG 60
Qy 19 ProTrpThrMetGluGluAspLeuLeuMetAsnTyrLeuAlaAsnHisGlyGluGly 38
Db 61 CCATGGACGATGGAAGAAGATTTCATCTCATCAATTATATCGCAATCATGGTGAAGGT 120
Qy 39 ValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db 121 GTTTGGAACTCTCGCCAAATTCGAGGACTAAACCGACCGGAAAGTTCGCCGCTC 180
Qy 59 ArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsnLeuThrProGluGluGln 78
Db 181 CGTGGCTGAACACTACCTCCGACCTGATGCGGCGGGAATATACACACAGAGACAG 240
Qy 79 LeuLeuLeuMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysLeuAlaLysHis 98
Db 241 CTCACCATCATGGAATTCATGCAAAATGGGAAATAGGTGTCACAAAATTCGAAAGCAT 300
Qy 99 LeuProGlyArgThrAspAsnGluLeuLysAsnTyrTrpArgThrArgLysGlnLysHis 118
Db 301 TTACCAGGAAGACCGCAATGAGATAAGAAATTTTGGAGGACTAAGATCCAGAAATAC 360
Qy 119 IleLysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluLeuAsnAspHis 138
Db 361 ATCATCAAGACGGAGAAACGACGCGTGGATCACAAGCTCCGAGTTTATAACCAT 420
Qy 139 GlnAlaSerThrSerHisVal---SerThrMetAlaGluProMetGluMetTyrSerPro 157
Db 421 CATCGCAACAGCGCCATGTCATGAATGATCTCAAGAAACCATGGATATGTATCTCCA 480
Qy 158 ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsn----- 175
Db 481 ACGACGTCGTATCAACATGCCAGCAATATTAAATCAGCAGCTTAATTATGGTAATTATGTG 540
Qy 176 ProAspGlnSerSerCys-----CysThrAsnAspAsnAsnAsnIleAsnTyrTrp 192
Db 541 CTGGAATCCGGTTCGATCATGATGCCATTATCTGTTGATCAATCCGAAACAACTATGG 600
Qy 193 SerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db 601 AGCGTCGATGATCTTGGCCCATGAATATATATATATGTAAT 642

RESULT 6

AY519589 681 bp mRNA linear PLN 07-FEB-2004
LOCUS Arabidopsis thaliana MYB transcription factor (At3g27810) mRNA,
DEFINITION complete cds.
ACCESSION AY519589
VERSION AY519589.1 GI:41619251
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 681)

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-659-869A-36 (1-206) x AY519589 (1-681)

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qu, L. and Gu, H.

The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide

Cloning and Expression Pattern Analysis

Unpublished

2 (bases 1 to 681)

Qu, L. and Gu, H.

Direct Submission

Submitted (07-JAN-2004) Life Sciences, National Laboratory of

Protein Engineering and Plant Genetic Engineering, Peking

University, Beijing 100871, China

Location/Qualifiers

1..681

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

1..681

/locus_tag="At3g27810"

1..681

/locus_tag="At3g27810"

/codon_start=1

/product="MYB transcription factor"

/protein_id="AA510059.1"

/db_xref="GI:41619252"

/translation="MEKRGSGSGSGSSAAEVRKPTWMEEDLILINVIANKGCV

WNSLAKAGLRKTKSRLRLWLYLRDPVRRGNITPEEQTLIMEHAKGNRRSKIAK

HLPRDNETKNFWRTRIQYIKOSDVTTSVSGSHSSEINDOAASTSHNVFCTQD

QAMETYSPTPTSYQHTNMEFNIGNYAAAAVTATVDYVPVMTVDQDGTGNTGMDIWS

SMHLNNGN"

Length: 681

Matches: 138

Conservative: 16

Mismatches: 44

Indels: 15

Gaps: 6

US-10-659-869A-36 (1-206) x AY519589 (1-681)

Qy 9 ThrSerGlnAspProGluValArgLysGlyProTrpThrMetGluGluAspLeuLeu 28

Db 40 TCATCAGCAGACGACGAAGTGAGAAAGGACCATGGACGATGGAAGAGATCTTATCTT 99

Qy 29 MetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaGly 48

Db 100 ATCACTATATCCCAACACCGCGGATGGTGTGGATTTCTCGCCAAATCTGCAGGT 159

Qy 49 LeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspVal 68

Db 160 CTAAACGAAACCGGAAAGTTCCGCTCGGTGCTGAATCTCTCCGCCGAGGTA 219

Qy 69 ArgArgGlyAsnIleThrProGluGluGlnLeuLeuMetGluLeuHisAlaLysTrp 88

Db 220 CGACGGGGAACATCATCTCAGAAAGCAACTTATCATCATGGAACCTTCATGTAAGTG 279

Qy 89 GlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluLys 108

Db 280 GGAACAGGTGGTTCGAAATTCGCAACATCTTCCAGGAAGACGACACGAGATCAA 339

Qy 109 AsnTyrTrpArgThrArgLysGlnLysHisIleLysGlnAlaGlu-----AsnPhe 125

Db 340 AATTTCTGGAGGCAAGAAATTCAAAATATACATCAAGCAATCGGATGTGAACACATCG 399

Qy 126 GlnGlnGlnSerSerAsnAsnSerGluLeuAsnAspHisGlnAlaSerThrSerHis--- 144

Db 400 TCCGTTGGATCTCATCATAGCTCAGAGATCAAGATCAAGCTGCAAGCAGCTGAGGCAT 459

Qy 145 -----ValSerThrMetAlaGluProMetGluMetTyrSer-----ProProCysTyr 160

Db 460 AATGTCTTTTGTACACAGATCAAGCGATGAGACTTATTTCTCTACACCGACATCATAT 519

Qy 161 GlnGlyMetLeuGluProPhe-----SerThrGlnPheProThrIleAsn 175

Db	520	CAACATACCAATATGGAATTCAACTATGTTAACTATTTCGGCCGCGGAGTGACGCGCAACC	579
Qy	176	ProAspGlnSerSerCysCysThr---AsnAspAsnAsnAsnIleAsnTyrTrpSerMet	194
Db	580	GTGGATTATCCAGTACCGATGACGTTGATGATCAAAACCGGTGAAGAACTATTGGGGCATG	639
Qy	195	GluAspSerTrp---SerMetGlnLeuLeuAsnGlyAsp	206
Db	640	GATGATATTGGTCATCAATGCAATTTATTGAATGGTAAT	678
RESULT 7			
LOCUS	AKI18439		
DEFINITION	Arabidopsis thaliana At3g27810 mRNA for putative myb family transcription factor (Atmyb3), complete cds, clone: RAF119-68-M03.		
ACCESSION	AKI18439		
VERSION	AKI18439.1	GI:26451910	
KEYWORDS	FLI cDNA; CAP trapper.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Arabidopsis thaliana full-length cDNA		
JOURNAL	Published Only in Database (2002)		
REFERENCE	2	(bases 1 to 1027)	
AUTHORS	Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: meeki@sc.riken.go.jp, URL: http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625, Fax: 81-45-503-9586)		
COMMENT	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.		
FEATURES	source		
	1..1027	Location/Qualifiers	
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="3"		
	/clone="RAF119-68-M03"		
	/ecotype="Columbia"		
	/note="common name: thale cress"		
gene	1..1027		
CDS	/gene="At3g27810/MGF10_23"		
	135..815		
	/gene="At3g27810/MGF10_23"		
	/codon_start=1		
	/product="putative myb family transcription factor Atmyb3"		
	/protein_id="BAC43048.1"		
	/db_xref="GI:26451911"		
	/translation="MEKRGSGSGSGSSAEEVKKGPWTWEEDLILINYIANHGDGV HNSLAKSLKRGSKRLWLYRPDVRGNITPEQLIMEHAKGNWSKIAK WLPGRITNEIKFNWRTRIQKTKOSDVTTSVGSHSSEINDQAASSTHNVFTQD QAMETYSTPTSYQHTNMEFNYSAAAVATATDYPVPMTVDQGTGNYNGMDIWS SMILNGN"		
ORIGIN			
Alignment Scores:			

	Pred. No.:	1..136-59	Length:	1027
	Score:	666.50	Matches:	138
	Percent Similarity:	72.30%	Conservative:	16
	Best Local Similarity:	64.79%	Mismatches:	44
	Query Match:	58.57%	Indels:	15
	DB:	8	Gaps:	6
	US-10-659-869A-36 (1-206) x AKI18439 (1-1027)			
Qy	9	ThrSerGlnAspProGluValArgLysGlyProTrpThrMetGluGluAspLeuIleLeu	28	
Db	174	TCATCAGCAGACAGAGAGTGAGAAAGGACCATGGAGCGATGGGAAGAGATCTTATCTT	233	
Qy	29	MetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaAlaGly	48	
Db	234	ATCAACTATATCGCCACACCGGGATGGTGTGGAAATCTCTCGCAATCTGCAGGT	293	
Qy	49	LeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspVal	68	
Db	294	CTAAACGAAACCGGAAAGTTGCGGCTCGGTGCTGAACATATCTCCGCCCGACGTA	353	
Qy	69	ArgArgGlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrp	88	
Db	354	CGACGGGAAACATCATCTCAGAAAGACCACTTATCATCTGGAACCTTCCTAAGTGG	413	
Qy	89	GlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLys	108	
Db	414	GGAAACAGGTGGTCGAAATTCGCCAAACATCTTCCAGGAAGACGACACGAGATCAA	473	
Qy	109	AsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGlu-----AsnPhe	125	
Db	474	AATTTCTGGAGGACAAGAATTCAAAAATACATCAAGCAATCGGATGTAACCAACATCG	533	
Qy	126	GlnGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHis---	144	
Db	534	TCCGTTGGATCTCATATAGCTCAGAGATCAACGATCAAGCTGCAGCACCTCGAGCCAT	593	
Qy	145	-----ValSerThrMetAlaGluProMetGluMetTyrSer-----ProProCysTyr	160	
Db	594	AATGCTCTTTGTACACAGATCAAGCATGAGAGACTTATCTCTCTACACCGACATCATAT	653	
Qy	161	GlnGlyMetLeuGluProPhe-----SerThrGlnPheProThrIleAsn	175	
Db	654	CAACATACCAATATGGAATTCACACTATGTAACATTTTCGGCCGCGGAGTACCGCAACC	713	
Qy	176	ProAspGlnSerSerCysCysThr---AsnAspAsnAsnAsnIleAsnTyrTrpSerMet	194	
Db	714	GTGATTTATCCAGTACCGATGACCGTTGATGATCAACCCGGTGAAACTATTGGGGCATG	773	
Qy	195	GluAspSerTrp---SerMetGlnLeuLeuAsnGlyAsp	206	
Db	774	GATGATTTGGTCATCAATGCAATTTATTGAATGGTAAT	812	
RESULT 8				
LOCUS	AR439844	Sequence 87 from patent US 6664446.	1033 bp	DNA linear PAT 20-FEB-2004
DEFINITION	AR439844			
ACCESSION	AR439844			
VERSION	AR439844.1	GI:42665809		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1033)			
AUTHORS	Heard, J., Broun, P., Riechmann, J. L., Keddie, J., Pineda, O., Adam, L., Samaha, R., Zhang, J., Yu, G.-L., Ratcliffe, O., Pilgrim, M., Jiang, C.-Z. and Reuber, L.			
TITLE	Transgenic plants comprising polynucleotides encoding transcription factors that confer disease tolerance			
JOURNAL	Patent: US 6664446-A 87 16-DEC-2003;			
FEATURES	Location/Qualifiers			
	1..1033			
	/organism="unknown"			


```

Db      634 CAACATACCAATATGGAATTTCAACTATGTTAACTATTTCGGCGCGCAGTGACGGCAACC 693
Qy      176 ProAspGlnSerSerCysCysThr---AsnAspAsnAsnIleAsnTyrTrpSerMet 194
Db      694 GTGGATTATTCAGTACCGATGACCGTTGATGATCAACACGGGTGAATAACTATTATGGGGCATG 753
Qy      195 GluAspSerTrp---SerMetGlnLeuLeuLeuAenGlyAsp 206
Db      754 GATGATATTTCGTATCAATGATGATTTATTAATGATGTAAT 792

RESULT 10
AY519582          621 bp mRNA linear PLN 07-FEB-2004
DEFINITION      Arabidopsis thaliana MYB transcription factor (At3g01530) mRNA,
SOURCE          complete cds.
ACCESSION      AY519582
VERSION        AY519582.1 GI:41619221
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
DEFINITION      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 621)
Qy,L. and Gu.H.
The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide
Cloning and Expression Pattern Analysis
Unpublished
2 (bases 1 to 621)
Qy,L. and Gu.H.
Submitted (07-JAN-2004) Life Sciences, National Laboratory of
Protein Engineering and Plant Genetic Engineering, Peking
University, Beijing 100871, China
Location/Qualifiers
1..621
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..621
/locus_tag="At3g01530"
1..621
/locus_tag="At3g01530"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AAS10052.1"
/db_xref="GI:41619221"
/translation="METTMMKKGRVKATITSQKEEGTVRKGPWTMBEDFILFNILN
HGEGLNWSVAKASGLKRTGKSLRWLNPDPVRRGNITEEQILLIQLHAKLGNRW
SKIAKHLPGRTDNETKNFWRTKIORHMKVSSBNMNMHOHCGNSQSGMTTQSSGK
AIDTAESFSQAKTTTFNVVEQQSNENYNNVEDLPVHLLNGDHVI"

ORIGIN
Alignment Scores:
Pred. No.:      3,546-52      Length:      621
Score:          592.00      Matches:    122
Percent Similarity: 65.60%      Conservative: 21
Best Local Similarity: 55.96%      Mismatches: 43
Query Match:     52.02%      Indels:     32
DB:              8          Gaps:       6

US-10-659-869A-36 (1-206) x AY519582 (1-621)

Qy      1 MetAspLysLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14
Db      13 ATGAAGAAGAAGGAGGAGTGAAGACGCAATACGTACAGAAAGAGAAGAGGAACA 72
Qy      15 ValAqLysGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsn 34
Db      73 GTGAGAAAAGGACCTTGGACTATGGAAGAGATTTCATCCCTTTAAATTACATCCTTAAT 132
Qy      35 HisGlyGluGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLys 54

```

```

Db      133 CATGCTGAAGGCTTTTGGAACTCTGTGCCTCCCAAAGCCTCTGCTCTAAACAGCTACTGGAATA 192
Qy      55 SerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThr 74
Db      193 AGTTGTCCGCTCCCGTGGCTGAATATCTCCGACAGATGTGCGCGGAGGAAACATAACC 252
Qy      75 ProGluGluGlnLeuLeuMetGlnLeuHisAlaLysTrpGlyAsnArgTrpSerLys 94
Db      253 GAAGAGACACAGCTTTTGCATCTTCAGCTTCATGCTTAAGCTTGGAAACAGGTGGTCGAAG 312
Qy      95 IleAlaLysHisLeuProGlyArgThrAspAsnGlnLysAsnTyrTrpArgThrArg 114
Db      313 ATTTCGAAGCATCTTCGGGAAGAACGACACACAGATAAAAGAACTTCTGGAGGACAAAC 372
Qy      115 IleGlnLysHisIleLysGlnAlaGlu-----AsnPheGlnGlnSerSer 130
Db      373 ATTCAGACACATGAAAGTGTCTCGAAATAATGATGAATCATCAATCATCTGTTTCG 432
Qy      131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
Db      433 GGAAACTCACAGACTCGGGGATGACGACGCAAGCAGCTCCGGCAAGCCATAGACACG 492
Qy      149 AlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuGluProPheSer 168
Db      493 GCTGAGAGCTTCTCT-----CAGGCGAAGACG 519
Qy      169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188
Db      520 ACGACGTTTAAATGGTGT-----GAACAACAGTCAAAAC 552
Qy      189 IleAsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db      553 GAGATTAATGGAACGTTGAAGATCTGTGGCCCTCCACTTGCTTAATGGTGAC 606

RESULT 11
BT005574          652 bp mRNA linear PLN 15-MAR-2003
LOCUS          Arabidopsis thaliana clone US0886 putative myb family transcription
DEFINITION      factor (At3g01530) mRNA, complete cds.
ACCESSION      BT005574
VERSION        BT005574.1 GI:28973338
KEYWORDS      FLI CDNA.
SOURCE          Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 652)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
2 (bases 1 to 652)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
Hayashizaki,Y. and Shinozaki,K.

```

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES

source

1..652 Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="U50886"

/ecotype="Columbia"

/note="This clone is in pUNI 51."

gene

1..652

/gene="At3g01530"

1..621

/gene="At3g01530"

/codon_start=1

/evidence=experimental

/product="putative myb family transcription factor"

/protein_id="AA063994.1"

/db_xref="GI:28973339"

/translation="METMKKGKVKATITQKEECTVRKGPWMEEDFILFNLIIN HGEGLNSVAKASGLKRTGKSCRLNLYLRPDVRGNITFEERQLLIQHLKLRHW SKIAHLPGTDNEIKTFWRKIQHMKVSVENMMNHQHCNSGSGMTTQSSOK AIDPAESFQAKTTFNVQSQSNWYVWEDLWVPLHLNGDHVI"

misc_difference

87

/gene="At3g01530"

/note="compared to genomic sequence"

622..652

/replace="t"

/gene="At3g01530"

ORIGIN

Alignment Scores:

Pred. No.: 3 75e-52 Length: 652
Score: 592.00 Matches: 122
Percent Similarity: 65.60% Conservative: 21
Best Local Similarity: 55.96% Mismatches: 43
Query Match: 52.02% Indels: 32
DB: 8 Gaps: 6

US-10-659-869A-36 (1-206) x BT005574 (1-652)

Qy 1 MetAspLysLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14
Db 13 ATGAAGAAGAAGGAGAGTGAACCGACATACGTCACAGAAAGAAAGAGGAACA 72
Qy 15 ValArgLysGlyProTrpThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsn 34
Db 73 GTGAGAAAGGACCATGAGTATGAGAGAGATTTCATCCTCTTTAATTACATCCTTAAT 132
Qy 35 HisGlyGluGlyValTrpAsnSerLeuAlaLysAlaGlyLeuLysAsnGlyLys 54
Db 133 CATGGTGAAGGTCCTTTGGAACTCTGCGCCAAAGCCCTCTGTGCTAAACGTAATCGGAAA 192
Qy 55 SerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsnIleThr 74
Db 193 AGTTGTGGCTCCGGTGGCTGAATCTTCGACCATGATGTCGCGGAGGAACATACC 252
Qy 75 ProGluGluGlnLeuLeuMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLys 94

Db 253 GAAGAAGAACAGCTTTTGATCATTCAGCTTCATGCTTAAGCTTGGAAACACAGGTGTCGAAG 312
Qy 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluLeuLysAsnTyrTrpArgTrpArg 114
Db 313 ATTGCGAAGCATTTTCGGGAAGAACCGACACGAGATAAAGAACTTCTGGAGGACAAAG 372
Qy 115 IleGlnLysHisLeuLysGlnAlaGlu-----AsnPheGlnGlnGlnSerSer 130
Db 373 ATTCAGAGACACATGAAGGTGTCATCGGAAATAATATGATGAATCATCAACATCATTTGTCG 432
Qy 131 AsnAsnSerGluLeuLeu-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 433 GGAACACTCAGAGACTCGGGAGTACGACCGACGCGAGCTCCGGCAAGGACATAGACACG 492
Qy 149 AlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuGluProPheSer 168
Db 493 GCTGAGAGCTTCTCT-----CAGGCGAAGACG 519
Qy 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysThrAsnAspAsnAsn 188
Db 520 ACGACGTTTAATGTGTG-----GAACCAACAGTCAAC 552
Qy 189 IleAsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db 553 GAGATTACTGGAACGTTGAGATCTGTGCGCCGTCACCTGCTTAATGTTGAC 606

RESULT 12

AK118091

LOCUS

DEFINITION

factor, complete cds, clone: RAFL19-34-G03.

ACCESSION

AK118091

VERSION

AK118091.1 GI:26451233

KEYWORDS

FLI CDNA; CAP trapper

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

1

Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Naruse, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

Arabidopsis thaliana full-length cDNA

Published Only in Database (2002)

2 (bases 1 to 894)

Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Naruse, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

Direct Submission

Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences

Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa

230-0045, Japan (E-mail: msekigsc.riken.go.jp)

URL: http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625,

Fax: 81-45-503-9586)

An Arabidopsis full-length cDNA library was constructed essentially

as reported previously (Seki et al. (1998) Plant J. 15:707-720;

Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et

al. (2001) Genomics 77:79-90) digested with BamHI and SalI.

This clone is in a modified pBluescript vector.

Please visit our web site (http://pfweb.gsc.riken.go.jp/) for

further details.

FEATURES

Location/Qualifiers

1..894

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAFL19-34-G03"

/ecotype="Columbia"

/note="common name: thale cress"

1..894

gene

```

CDS
/gene="At3g01530/F4P13_8"
109..729
/codon_start=1
/product="putative transcription factor"
/db_xref="GI:26451234"
HGEGLNWSAVAKSLKRTGKSCRLLWNLRLPDVRGNITTEEOQLLIQLHAKLGNRW
SKIAKHLPGRTDNEIKNFWTKIQRMKVSSENMMNHQHCNSQSGMTTQSSGK
AIDTAESFSQAKTTFNFVVEQOSNENYNNVEDLWPFVHLLNGDHVI"

ORIGIN
Alignment Scores:
Pred. No.: 5,46e-52 Length: 894
Score: 592.00 Matches: 122
Percent Similarity: 65.60% Conservatives: 21
Best Local Similarity: 55.96% Mismatches: 43
Query Match: 52.02% Indels: 32
DB: 8 Gaps: 6

US-10-659-869A-36 (1-206) x AK118091 (1-894)

Qy 1 MetAspLysLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14
Db 121 ATGAAGAAGAAAGGAGAGTGAAGCGCAATAACGTCACAGAAAGAAAGAGGAACA 180
Qy 15 ValArgLysGlyProTrrThrMetGluAspLeuIleLeuMetAsnTyrIleAlaAsn 34
Db 181 GTGAAAGAGGACCATGACTATGGAAGAAGATTTTCCTCTTTAAATTAGCTCTTAAT 240
Qy 35 HisGlyGluGlyValTrrAsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLys 54
Db 241 CATGTGAGGCTTTTGGACTCTGTCCCAAGGCTCTGTCTTAAACGCTACTGGAAA 300
Qy 55 SerCysArgLeuArgTrrLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThr 74
Db 301 AGTTCTCGCTCCGGTGGCTGAACATATCTCCGACCAGATGTGCGGAGGAAACATAACC 360
Qy 75 ProGluGluGlnLeuIleMetGluLeuHisAlaLysTrrGlyAsnArgTrrSerLys 94
Db 361 GAAGAAGAACACGCTTTTATCATCTTCACTTAAGCTTGGAAACAGGTGGTCGAAG 420
Qy 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrrArgThrArg 114
Db 421 ATTGCGAGCATCTTCGGGAAGACGCGAACACGAGATTAAGAACTTCTGGAGGACAAAG 480
Qy 115 IleGlnLysHisIleLysGlnAlaGlu-----AsnPheGlnGlnGlnSerSer 130
Db 481 ATTCAGAGACACATGAAGTGTCTCGGAAATATGATGAATCATCAACATCATTTGTTTCG 540
Qy 131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 541 GGAAACTCACAGAGTCGGGGATACGACGCAAGCAGCTCCGGCAAGCCCATAGACACG 600
Qy 149 AlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuGluProPheSer 168
Db 601 GCTGAGAGCTTCTCT-----CAGCGGAGACG 627
Qy 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTrrAsnAspAsnAsn 188
Db 628 ACGAGCTTTAATGTGGT-----GAACAACAGTCAAAAC 660
Qy 189 IleAsnTyrTrrSerMetGluAspSerTrrSerMetGlnLeuLeuAsnGlyAsp 206
Db 661 GAGAATTACTGAAGCTTGAAGATCTGTGCCCGCTCCACTTGTCTTAATGGTGAC 714

RESULT 13
AY088761
LOCUS AY088761 1127 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 94595 mRNA, complete sequence.
ACCESSION AY088761
VERSION AY088761.1 GI:21407535

KEYWORDS
SOURCE ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 1127)
Haas,B.J., Volkovskiy,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
JOURNAL MEDLINE 22088475
PUBMED 12093376
REFERENCE
2 (bases 1 to 1127)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
REFERENCE
3 (bases 1 to 1127)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genset carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
location/Qualifiers
1..1127
/orf1.gene="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="94595"
274..894
/codon_start=1
/product="putative transcription factor"
/protein_id="AAM67076.1"
/db_xref="GI:21618026"
/translation="METMTKKKGRVKATITTSQKEEGTVRKGPWTMEEDFLFNYILN
HGEGLNWSAVAKSLKRTGKSCRLLWNLRLPDVRGNITTEEOQLLIQLHAKLGNRW
SKIAKHLPGRTDNEIKNFWTKIQRMKVSSENMMNHQHCNSQSGMTTQSSGK
AIDTAESFSQAKTTFNFVVEQOSNENYNNVEDLWPFVHLLNGDHVI"

ORIGIN
Alignment Scores:
Pred. No.: 3,85e-51 Length: 1127
Score: 585.00 Matches: 121
Percent Similarity: 65.14% Conservatives: 21
Best Local Similarity: 55.50% Mismatches: 44
Query Match: 51.41% Indels: 32
DB: 8 Gaps: 6

US-10-659-869A-36 (1-206) x AY088761 (1-1127)

Qy 1 MetAspLysLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14
Db 286 ATGAAGAAGAAAGGAGAGTGAAGCGCAATAACGTCACAGAAAGAAAGAGGAACA 345
Qy 15 ValArgLysGlyProTrrThrMetGluAspLeuIleLeuMetAsnTyrIleAlaAsn 34
```



```

Db 346 GTGAGAAAGACCTTGGACTATCGAAGAAATTCATCTCTTTAATTACATCTTAAT 405
Qy 35 HisGlyGluGlyValTTPAsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLys 54
Db 406 CATGGTAGGCTCTTGGAACTCTGTGCCAAAGCTCTGTCTTAAACGCTACTGGAAA 465
Qy 55 SerCysArgLeuArgTTPLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThr 74
Db 466 AGTTGTCCGCTCCGCTGGCTGAATCTCCGACAGATGTGCGCGAGGAGCAATAACC 525
Qy 75 ProGluGluGlnLeuLeuLeuMetGluLeuHisAlaLysTTPGlyAsnArgTTPSerLys 94
Db 526 ACAGAAGAAGACGCTTTTGCATCATTCAGCTTATAGCTTGGAAAACAGGTGTCGAAG 585
Qy 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluLysAsnTyrTTPArgThrArg 114
Db 586 ATTCCGAAGCATCTTCGGGAAGAACGACCAACAGATTAAGAATCTCTGGAGACAAAG 645
Qy 115 IleGlnLysHisIleLysGlnAlaGlu-----AenPheGlnGlnSerSer 130
Db 646 ATTCAGAGACATGAAAGTGTATCGGAAATATGATGAATCATCAACATCATTTGTCG 705
Qy 131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 706 GGAAGACTCACAGACTCGGGGATGACGACGCAAGGACGCTCCGGCAAGCCATAGACACG 765
Qy 149 AlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuGluProPheSer 168
Db 766 GCTGAGAGCTTCTCT-----CAGGCGAAGACG 792
Qy 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysThrAsnAspAsnAsn 188
Db 793 ACGAGTTTAAATGGTG-----GAACAACAGCTCAAC 825
Qy 189 IleAsnTyrTTPSerMetGluAspSerTTPSerMetGlnLeuLeuAsnGlyAsp 206
Db 826 GAGAATTACTGGAACGTTGGAGATCTGTGGCCCGTCCACTTGTCTTAATGGTGAC 879

RESULT 14
LOCUS AB058642 872 bp mRNA linear PLN 04-APR-2001
DEFINITION Liliium hybrid division I LhMyb mRNA, complete cds.
ACCESSION AB058642
VERSION AB058642.1 GI:13537529
KEYWORDS
SOURCE Liliium hybrid division I
ORGANISM Liliium hybrid division I
REFERENCE
AUTHORS Nakatsuka, A., Izumi, Y. and Yamagishi, M.
TITLE Isolation and characterization of the genes related to anthocyanin biosynthesis in Asiatic hybrid lily
JOURNAL Unpublished
AUTHORS Nakatsuka, A., Izumi, Y. and Yamagishi, M.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2001) Akira Nakatsuka, Shimane University, Faculty of Life and Environmental Sciences; Nisikawatsu 1060, Matsue, Shimane 690-8504, Japan
(E-mail:nakira@life.shimane-u.ac.jp, Tel:81-852-32-6502)
Location/Qualifiers
FEATURES
source 1..872
/organism="Liliium hybrid division I"
/mol_type="mRNA"
/cultivar="Montreux"
/db_xref="taxon:156532"
/tissue_type="tepals"
1..872
/gene="LhMyb"
80..634
gene
CDS

```

```

/gene="LhMyb"
/codon_start=1
/protein_id="BAB40790.1"
/db_xref="GI:13537530"
/translation="MDKVTGNTEDVEVRKGPWTMEEDLILINIANHGDGVNMTLARS
AGLXPTGKSRLEWLYLAPDVRGNITPEEQILLIMELHDMRGNRWSKIAKEUPGRD
NEIKFWIRTVKQKSKNGESSEQISMLSDETSAMQMEDVDVGSQTSYQGGQTSNNL
DAFEIPEFGDGNLLSVEELWAMQY"
ORIGIN
Alignment Scores: 1.03e-49 Length: 872
Pred. No.: 570.00 Matches: 118
Score: 570.00 Conservations: 17
Percent Similarity: 64.29% Mismatches: 39
Best Local Similarity: 56.19% Indels: 36
Query Match: 50.09% Gaps: 5
DB: 8
US-10-659-869A-36 (1-206) x AB058642 (1-872)
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTTP 20
Db 80 ATGCACAGGAAA-----GTAGGAGACTAATGAAGATGTTTCAGGTGAGGAAAGGACCATGG 133
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTTP 40
Db 134 ACCATGGAAGAGGATCTCATCTCACTACATAGCCCAACCATGCGCGAGGCTCTGG 193
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlnLysSerCysArgLeuArgTTP 60
Db 194 AACACGCTGGCTCGATCGGCCGAGCTGAAGAGAGCTGGAAGAGAGCTCAGGCTTCGTGG 253
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 254 CTGAATTTACCTGAGCGCTGATGTTCTGCGGGTAACATCACACAGAGGAGCAGCTTCTG 313
Qy 81 IleMetGluLeuHisAlaLysTTPGlyAsnArgTTPSerLysLysIleAlaLysHisLeuPro 100
Db 314 ATCATGGAACCTTCACGACAGATGGGTGAACAGATGGTCGAAAAATTGCAAGAGGAGCTCCA 373
Qy 101 GlyArgThrAspAsnGluLysAsnTyrTTPArgThrArgIleGlnLysHisIleLys 120
Db 374 GGACGAACAGATTAACGAATCAAGAACTCTTGAGGAGACAGAGCTCCAGAG----- 424
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 425 -----AAATCCAAGATGCGCAATCATCTGAGGCGCAAAATC 460
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGlu----- 153
Db 461 TCCATGCTTTCCGATGAAACTAGTCCCATGCAATGGAAGATGTCGATGTTGGTGTGTGTCG 520
Qy 154 -----MetTyrSerProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhe 171
Db 521 CAGACGAGTACCATCAAGCCGAGACTTCGAACACCTTGACGCAATTTGAAATTCCTTTC 580
Qy 172 ProThrIleAsnProAspGlnSerSerCysThrAsnAspAsnAsnIleAsnTyr 191
Db 581 -----GAGCTCGCGAT-----AATCTC 598
Qy 192 TrpSerMetGluAspSerTTPSerMetGln 201
Db 599 CTGTCCGTCGAGGAACCTGTGGCTATGCAG 628

RESULT 15
LOCUS AF485900 649 bp mRNA linear PLN 27-AUG-2003
DEFINITION Dendrobium sp. XMW-2002-9 MYB9 (MYB9) mRNA, complete cds.
ACCESSION AF485900
VERSION AF485900.1 GI:28628962
KEYWORDS
SOURCE Dendrobium sp. XMW-2002-9
ORGANISM Dendrobium sp. XMW-2002-9

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae; Epidendroideae; higher Epidendroideae; Dendrobieae; Dendrobiinae; Dendrobium.

REFERENCE 1 (bases 1 to 649)
AUTHORS Wu,X.-M., Lim,S.-H. and Yang,W.-C.
TITLE Characterization, expression and phylogenetic study of R2R3-MYB genes in orchid

JOURNAL Plant Mol. Biol. 51 (6), 959-972 (2003)
MEDLINE 22661209
PUBMED 12777054

REFERENCE 2 (bases 1 to 649)
AUTHORS Wu,X.-M., Lim,S.-H. and Yang,W.-C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Laboratory of Plant Reproduction & Cotton Biotechnology, Institute of Molecular Agrobiolology, 1 Research Link S117604, Singapore

FEATURES Location/Qualifiers

source 1..649

organism="Dendrobium sp. XMW-2002-9"

mol_type="mRNA"

/db_xref="taxon:217455"

gene 1..649

/gene="MYB9"

CDS 21..476

/gene="MYB9"

/note="MYB"

/codon_start=1

/product="MYB9"

/protein_id="AAO49418.1"

/db_xref="GI:28628963"

/translation="WEEDGILIN IYANHGEVWNNAKLSAGLKRGTGKSCRRLWNLNLR PDVPRGNITPEQLLIMELHARWGNRSKIARQLPGRDNEIKNWRTRIQKVKNGD AVDYSSQMIIDEACTSSIEEGMAQQNPYPVHPNTNPEGYRLQTRVN"

ORIGIN

Alignment Scores:

Pred. No.: 5,58e-49 Length: 649
Score: 561.50 Matches: 114
Percent Similarity: 71.51% Conservative: 19
Best Local Similarity: 61.29% Mismatches: 36
Query Match: 49.34% Indels: 17
DB: 8 Gaps: 2

US-10-659-869A-36 (1-206) x AF485900 (1-649)

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGlyValTrp 40
Db |||||
18 ACTATGGAAGAGCGGATCCTCATAACTATATAGCCACCATCGAGAGGAGTATGG 77
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db |||||
78 AATAATCTTGCCAAATCTGCAGGTTTGAGAGGACGGGGAAGAGCTGCCGGCTTCGATGG 137
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db |||||
138 CTAAACTATCTGAGGCCACAGCTGAGAGGGGGAACATACACCGAGGAGCAGCTGCTG 197
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db |||||
198 ATCATGGAGTTGCATGCCAGCTGGGGAAACAGGTGTCAAAAATTGCGAGGCAACTCCA 257
Qy 101 GlyArgThrAspAsnGluLeuLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db |||||
258 GGAAGACACTGAATGAGATCAAGAACTACTCTGAGGACTAGAAATTCAAAAGAGGTTAAG 317
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db |||||
318 AATGGGATGCAGTCGATTAC-----TCGAGCCAAATGATCATTCATGATGAGGCC 365
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
Db |||||
366 TGCACAAGCCATACAAGTAGCTAGTAAGAAGGAATGGCTCAGCAAAATTTATCTCTTCAC 425

Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db -----CCTAATACCAATCTTGAAAGGTATCGT 452
Qy 181 CysCysThrAsnAspAsnAsn-AsnIleAsnTyrTrpSerMetGluAsp-SerTrpSerM 200
Db |||||
453 CTTCAAAACAACACACGGGTGAATTGACGAGAACTTTTGGAGTGTGGAGGATTTCTTGGTCCA 512
Qy 200 etGlnLeuLeuAsn 204
Db |||||
513 CCCAATCTGCCAAT 526

Search completed: September 28, 2005, 14:40:38
Job time : 4146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2005, 13:19:02 ; Search time 527 Seconds
(without alignments)
2313.979 Million cell updates/sec

Title: US-10-659-869A-36
Perfect score: 1138
Sequence: 1 MDKKQCKTSQDPVRKGPW.....NNINYSMEDSWSMQLLNGD 206

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10659869/runat_26092005_143444_11364/app_query.fasta_1.391
-DB=N Geneseq_16Dec04 -QWMT=rascp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10659869 @CGN 1 1 470 @runat_26092005_143444_11364 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1138	100.0	863 8	ACD07386 Soybean M
2	1138	100.0	863 12	ADJ77736
3	1092	96.0	968 12	ADJ43074 Plant tra
4	1092	96.0	968 12	ADO03218 Soybean o
5	1062	93.3	988 8	ACD07383 Soybean M

6	1062	93.3	988	12	ADJ77730	Adj77730 cDNA enco
7	889.5	78.2	988	12	ADJ43073	Adi43073 Plant tra
8	889.5	78.2	988	12	ADO03217	Ado03217 Soybean o
9	874.5	76.8	805	8	ACD07387	AcD07387 Soybean M
10	874.5	76.8	805	12	ADJ77738	Adj77738 cDNA enco
11	829.5	72.9	910	8	ACD07385	AcD07385 Soybean M
12	829.5	72.9	910	12	ADJ77734	Adj77734 cDNA enco
13	810	71.2	969	12	ADJ43075	Adi43075 Plant tra
14	810	71.2	969	12	ADO03219	Ado03219 Soybean o
15	760	66.8	681	10	ADD17714	Add17714 DNA (Seq)
16	760	66.8	681	10	ADK56061	AdK56061 Plant DNA
17	759.5	66.7	862	3	AAC57336	Aac57336 Eucalyptu
18	724	63.6	681	12	ADJ10801	Adj10801 Recombina
19	671	59.0	910	3	AAC40587	Aac40587 Arabidops
20	667	58.6	908	3	AAC48506	Aac48506 Arabidops
21	663	58.3	731	12	ADJ43518	Adi43518 Plant tra
22	663	58.3	731	12	ADO03344	Ado03344 Thalecres
23	663	58.3	731	12	ADO01760	Ado01760 Thalecres
24	653.5	57.4	1033	9	ACD98402	AcD98402 A. thalia
25	653.5	57.4	1033	10	ADD55727	Add55727 Thalecres
26	653.5	57.4	1033	10	ADD30419	Add30419 Plant yie
27	653.5	57.4	1033	12	ADJ43506	Adi43506 Plant tra
28	653.5	57.4	1033	12	ADO03332	Ado03332 Thalecres
29	638.5	56.1	462	10	ADJ17692	Adj17692 DNA (Seq)
30	638.5	56.1	462	10	ADK59524	AdK59524 Plant DNA
31	625.5	55.0	907	3	AAC57337	Aac57337 Eucalyptu
32	592	52.0	1011	10	ADP30197	Adp30197 Plant yie
33	592	52.0	1011	10	ADK37062	Adk37062 Plant yie
34	592	52.0	1011	10	ADK37082	Adk37082 Plant yie
35	592	52.0	1011	10	ADK31530	Adk31530 Plant yie
36	592	52.0	1011	12	ADJ41794	Adj41794 Plant tra
37	592	52.0	1011	12	ADJ41794	Adj41794 Plant tra
38	592	52.0	1011	12	ADJ41794	Adj41794 Plant tra
39	591	51.9	530	8	ACD07384	AcD07384 Soybean M
40	591	51.9	530	12	ADJ77732	Adj77732 cDNA enco
41	587.5	51.6	392	12	ADQ06964	Adq06964 Soybean t
42	585	51.4	621	3	AAC41714	Aac41714 Arabidops
43	585	51.4	1127	3	AAC37953	Aac37953 Arabidops
44	540	47.5	673	3	AAC56197	Aac56197 Eucalyptu
45	539	47.4	1137	4	AAD05784	Aad05784 Arabidops

ALIGNMENTS

RESULT 1
ACD07386
ID ACD07386 standard; cDNA; 863 BP.
XX
AC ACD07386;
XX
DT 07-AUG-2003 (first entry)
XX
DE Soybean Myb-related transcription factor #5 cDNA.
XX
KW Soyben; ss; Myb-related transcription factor; crop improvement;
XX starch composition; protein composition; plant; gene; grain quality;
XX disease resistance; pest resistance; oil composition.
XX Glycine max.
XX
FN US2003024007-A1.
XX
PD 30-JAN-2003.
XX
PF 14-DEC-2001; 2001US-00021811.
XX
PR 02-DEC-1998; 98US-0110609P.
PR 01-DEC-1999; 99US-00452244.
XX
PA (CAHO/) CAHOON R E.
XX (ODEL/) ODELL J T.
PI Cahoon RE, Odell JT;

XX WPI; 2003-456317/43.
 DR P-PSDB; AB001759.
 XX New isolated polynucleotide encoding a Myb-related transcription factor,
 PT useful in plant molecular biology, in particular for the improvement of
 PT crop plants for a variety of traits, including disease and pest
 PT resistance.
 XX Claim 2; Page 36; 61pp; English.
 XX The invention relates to an isolated polynucleotide encoding a Myb-
 CC related transcription factor polypeptide. The methods and compositions of
 CC the present invention are useful in plant molecular biology, in
 CC particular for the improvement of crop plants for a variety of traits,
 CC including disease and pest resistance and grain quality improvements such
 CC as oil, starch or protein composition. The present sequence represents
 CC a cDNA encoding a soybean Myb-related transcription factor
 XX
 SQ Sequence 863 BP; 306 A; 149 C; 180 G; 225 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 1.44e-115 Length: 863
 Score: 1138.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-10-659-869A-36 (1-206) x ACD07386 (1-863)
 QY 1 MetAspLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
 Db 29 ATGGATAAAAAACACAGGTGTAGACGCTCTCAAGATCTCTGAAGTGAAGAAAGGCGCTTGG 88
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
 Db 89 ACAATGGAAGAAGACTTGTATCTTGATGAACATATATTGCAAAATCATGTGGGAAGGTGTTGG 148
 QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
 Db 149 AACTCTTTGGCCAAAGCTGCTGGTCTCAACGTAACGGAAGAGTTCGGCGCTAAGGTGG 208
 QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 209 CTAAATTAACCTCCGTCCTGATGTTAGAGAGGGGAATATATACCCGAGGAAACAACATTTTG 268
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
 Db 269 ATTATGGAGCTCCACGCAAGTGGGGAACAGGTGCTCCAAATTTGCCAAGCATCTACCT 328
 QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
 Db 329 GGAAGGACTGTAATAGATCAAGAACATATTGGAGGACAAGATCCAGAACACATCAAG 388
 QY 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db 389 CAAGCTGAGAACTTTCAGCAACAGAGTAGTAATAATTTCTGATATAATGATCACCAGCT 448
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
 Db 449 AGCACTAGCATGTTTCCACCATGCTGAGCCCATGGAGATGATATCTCCACCCCTGTTAT 508
 QY 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
 Db 509 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTACAAATTAATTCCTGATCAATCCAGT 568
 QY 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
 Db 569 TGTTGTACCAATGACACACACACATTAATTTGAGCATGGAGATAGCTGGTCAATG 628
 QY 201 GlnLeuLeuAsnGlyAsp 206
 |||||||

Db 629 CAATTACTGAACGGTGAT 646
 RESULT 2
 ADJ77736
 ID ADJ77736 standard; cDNA; 863 BP.
 XX AC ADJ77736;
 XX 20-MAY-2004 (first entry)
 XX cDNA encoding soybean Myb-related transcription factor #5.
 XX Plant; Myb-related transcription factor; disease resistance; soybean;
 KW gene; ss.
 XX Glycine max.
 OS
 XX US2004040057-A1.
 FN
 XX 26-FEB-2004.
 PD
 XX 11-SEP-2003; 2003US-00659869.
 PF
 XX 02-DEC-1998; 98US-0110609P.
 PR
 XX 01-DEC-1999; 99US-00452244.
 PR
 XX 14-DEC-2001; 2001US-00021811.
 XX
 PA (CAHO/) CAHOON R E.
 PA (FANG/) FANG Y.
 PA (ODEL/) ODELL J T.
 PA (WENG/) WENG Z.
 XX
 XX Cahoon RE, Fang Y, Odell JT, Weng Z;
 DR WPI; 2004-238121/22.
 DR P-PSDB; ADJ77737.
 XX
 PT Plant Myb transcription factor homologs and their polynucleotide
 PT sequences useful in plant genomic research e.g. identifying nucleotide
 PT fragments that encode a Myb-related transcription factor polypeptide.
 XX
 PS Claim 2; SEQ ID NO 35; 62pp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding plant Myb-related transcription factors and the
 CC encoding polypeptide sequences. Also disclosed is a chimeric gene
 CC encoding all or part of a Myb-related transcription factor linked to
 CC suitable regulatory sequences, and a host cell transformed with the
 CC chimeric gene and able to express the chimeric gene resulting in the
 CC production of altered levels of the Myb-related transcription factor in
 CC the transformed cell. The polynucleotide sequences of the invention may
 CC be used to probe a cDNA or genomic library to identify nucleic acid
 CC fragments encoding a Myb-related transcription factor polypeptide and
 CC also for research purposes including generating plants with increased
 CC resistance to disease, etc. The present sequence encodes a plant Myb-
 CC related transcription factor.
 XX
 SQ Sequence 863 BP; 306 A; 149 C; 180 G; 225 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 1.44e-115 Length: 863
 Score: 1138.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 US-10-659-869A-36 (1-206) x ADJ77736 (1-863)
 QY 1 MetAspLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
 Db 29 ATGGATAAAAAACACAGGTGTAGACGCTCTCAAGATCTCTGAAGTGAAGAAAGGCGCTTGG 88

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
 Db 89 ACAATGGAAGAGACTTTCATCTTCATGAACATATATTCAAATCATATGCGGAAGGTGTTGG 148
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysAtgLeuArgTrp 60
 Db 149 AACTCTTTGGCAAGCTGCTGGTCTCAACGTAACGGAAAGAGTTCCGGCTAAGGTGG 208
 Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 209 CTAAATACCTCCCTCGTCTGATGTTAGAGAGGGAATATACACCCGAGGACAACTTTTG 268
 Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
 Db 269 ATTTAGAGCTCCACGCAAGTGGGAAACAGGTGGTCCAAATTTGCCAAGCATCTACCT 328
 Qy 101 GlyArgThrAspAsnGluLysLeuLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
 Db 329 GGAAGGACTGATATGAGATCAAGAACTATTGGAGGACAGGATCCAGAGCATCAAG 388
 Qy 121 GlnAlaGluAsnPheGlnGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db 389 CAAGCTGAGAAGCTTTCAGCAACAGAGTAGTAATAATTCTGAGATAAATGATCAACAAGCT 448
 Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
 Db 449 AGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTATTTCCACCCCTGTAT 508
 Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
 Db 509 CAAGAAATGTTAGAGCCATTTTCAACTCAGTTTCCTTCAATTAATCTGATCAATCCAGT 568
 Qy 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
 Db 569 TGTGTACCAATGACACACACCAATTAATTTGGAGCATGGAGGATAGTGTGTCATG 628
 Qy 201 GlnLeuLeuAsnGlyAsp 206
 Db 629 CAATTACTGAACGGTGAT 646
 RESULT 3
 ID AD143074 standard; DNA; 968 BP.
 AC XX
 CC AD143074;
 DT 22-APR-2004 (first entry)
 DE Plant transcription factor polynucleotide #1009.
 DE XX
 KW transgenic; plant; enhanced tolerance to abiotic stress;
 KW glyphosphate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.
 XX Glycine max.
 OS XX
 XX US2004019927-A1.
 XX 29-JAN-2004.
 XX 25-FEB-2003; 2003US-00374780.
 XX 18-APR-2001; 2001US-00837944.
 XX (SHERMAN B K.
 XX (RIECHMANN J L.
 XX (JIAN) JIANG C.

PA (HEAR/) HEARD J E.
 PA (HAAK/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (PILG/) PILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V,
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX WPI; 2004-132245/13.
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 PS Claim 1; SEQ ID NO 1537; 435pp; English.
 XX The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produced a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure; change in stem bifurcations; altered branching
 CC pattern; reduced apical dominance; reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins; or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factors isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.
 XX SQ Sequence 968 BP; 308 A; 197 C; 195 G; 268 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.99e-110 Length: 968
 Score: 1092.00 Matches: 198
 Percent Similarity: 97.09% Conservative: 2
 Best Local Similarity: 96.12% Mismatches: 6
 Query Match: 95.96% Indels: 0
 DB: 12 Gaps: 0
 US-10-659-869A-36 (1-206) x AD143074 (1-968)
 Qy 1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
 Db 72 ATGGATAAAAAACAACACTGTGCAACACGCTCTCAAGATCCTCAAGTAGAGAAAGGCGCTGG 131
 Qy 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
 Db 132 ACAATGGAAGAGACTTTCATCTTCATCACTATATTCGAATCATGGGAGGTGTTGG 191
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
 Db 192 AACTCTTTGGCAAGCTGCTGGTCTCAACGTAACCGAAGAGTTCCGGCTAAGGTGG 251
 Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80

Db 252 CTAACCTACCTCCGTCCTGATGTTAGAGAGGGGAATATTACACCCGAGGACAACTTTTG 311
Qy 1leMetGluLeuHisAlaLysTyrGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 312 ATCATGGAGCTTCACGAAAGTGGGAAACAGGTGGTCCAAAATGCCAAGCATCTACCT 371
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 372 GGAAGGACTGATATGAGATCAAGAACTATTGGAGGACAAGGATCCAGAAGCACATCAAG 431
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 432 CAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATCTGAGATAAAATGATCCCAAGCT 491
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 492 TGCCTACCTGATGTCACCATGTCGCGAGCCCATGGAGACTATTCTCCACCTGTTAT 551
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 552 CAAGGAATGTTAGAGCCATTTTCACTCAGTTCCTACAAATTAATCTCGATCAATCCAGT 611
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db 612 TGTGTACCAATGACAAACAACATTAATTAATTTGGAGCATGGAGATAGCTGGTCAATG 671
Qy 201 GlnLeuLeuAsnGlyAsp 206
Db 672 CAATTACTGAACGGGTGAT 689

RESULT 4
AD003218
ID AD003218 standard; cDNA; 968 BP.
XX AC AD003218;
XX DT
XX DE
XX DE
XX DE
KW Soybean orthologue of Thalecress transcription factor, cDNA #270.
KW Soybean; transcription factor; ss; gene; plant; transgenic;
KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
KW phosphate limitation; potassium limitation; nitrogen limitation;
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
KW flowering; inflorescence architectural change;
KW meristem cell differentiation; phyllotaxy; apical dominance;
KW trichome development; seed development; premature senescence;
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
KW seed morphology; secondary metabolism; light response; shade avoidance.
XX OS Glycine max.
XX PN US2004045049-A1.
XX PD
XX PD
XX PF 10-APR-2003; 2003US-00412699.
XX XX
PR 13-SEP-1999; 99US-00394519.
PR 21-JAN-2000; 2000US-00489376.
PR 17-FEB-2000; 2000US-00506720.
PR 22-MAR-2000; 2000US-00532591.
PR 22-MAR-2000; 2000US-00533029.
PR 22-MAR-2000; 2000US-00533030.
PR 22-MAR-2000; 2000US-00533392.
PR 22-MAR-2000; 2000US-00533648.
PR 06-APR-2000; 2000WO-US009448.
PR 16-NOV-2000; 2000US-00713994.
PR 27-MAR-2001; 2001US-00819142.
PR 17-APR-2001; 2001US-00837444.
PR 30-JAN-2002; 2002US-00958131.
PR 14-JUN-2002; 2002US-00171468.
PR 09-AUG-2002; 2002US-00225066.

PR 09-AUG-2002; 2002US-00225067.
PR 09-AUG-2002; 2002US-00225067.
PR 17-DEC-2002; 2002US-0434166P.
PR 25-FEB-2003; 2003US-00374780.
XX (ZHAN/) ZHANG J.
PA (FROM/) FROMM M E.
PA (HEAR/) HEARD J E.
PA (RIEC/) RIECHMANN J L.
PA (ADAM/) ADAM L J.
PA (BROU/) BROUN P E.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J S.
PA (YUGG/) YU G.
PA (JIAN/) JIANG C.
PA (SAMA/) SAMAHA R S.
PA (PILG/) PILGRIM M L.
PA (CREE/) CREELMAN R A.
PA (DUBE/) DUBELL A N.
PA (RATC/) RATCLIFFE O.
PA (KIMI/) KUMIMOTO R.
PA (SHER/) SHERMAN B K.
XX
PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI Sherman BK;
XX WPI; 2004-225755/21.
DR
XX New transgenic plant, useful in developing phenotypes with altered or improved characteristics or traits.
PT Claim 1; SEQ ID NO 1632; 213pp; English.
XX
XX The invention relates to a transgenic plant comprises a recombinant polynucleotide having a polynucleotide sequence or its complementary sequence comprising a sequence encoding a polypeptide, that initiates transcription (i.e. a transcription factor) from Arabidopsis, Soybean, Rice, Rape or Corn, comprising any of the sequences appearing as AD001588 -AD003527 or AD003530-AD003559. Also included are using a transgenic plant to grow a progeny plant, an expression cassette (comprising a constitutive, inducible or tissue-specific promoter and a recombinant polynucleotide described above), a host cell comprising the expression cassette, producing a modified plant having a modified trait, identifying a factor that is modulated by or interacts with a polypeptide encoded by the polynucleotide sequence and identifying at least one downstream polynucleotide sequence that is subject to a regulatory effect of any of the polypeptides encoded by the polynucleotide described above. The transgenic plant is useful for producing a plant that has an altered trait e.g. an enhanced tolerance to abiotic stress (increased tolerance to chilling, germination in cold conditions, freezing tolerance, tolerance to heat, tolerance to drought, tolerance to osmotic stress, tolerance to salt, tolerance to phosphate limitation, tolerance to potassium limitation, decreased sensitivity to nitrogen limitation), altered hormone sensitivity, reduced resistance to abscisic acid, an altered response to ethylene, disease resistance, altered susceptibility to Botrytis, altered susceptibility to Fusarium, altered susceptibility to Erysiphe, altered susceptibility to Pseudomonas syringae, altered susceptibility to Sclerotinia, altered sugar sensing, improved seed germination and seedling vigor, early flowering, late flowering, extended period of flowering, an inflorescence architectural change, a change in stem bifurcations, a lack of a shoot meristem, reduced meristem cell differentiation, altered phyllotaxy, altered branching pattern, reduced apical dominance, reduced trichome density, ectopic trichome development, altered trichome development, altered seed morphology, increased root growth, increased root hairs, altered stem development, delayed cell proliferation/cell differentiation, premature senescence, delayed senescence, lethality, increased necrosis, an increase in seedling or plant size, decreased plant size, a change in leaf morphology, increased altered leaf development, increased leaf size and mass, glossy leaves, leaf cell expansion, change in seed morphology, altered seed coloration,

CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenly lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenly lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence encodes an
 CC orthologue of a thalecress transcription factor isolated from Soybean.
 XX
 SQ Sequence 968 BP; 308 A; 197 C; 195 G; 268 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.99e-110 Length: 968
 Score: 1092.00 Matches: 198
 Percent Similarity: 97.09% Conservative: 2
 Best Local Similarity: 96.12% Mismatches: 6
 Query Match: 95.96% Indels: 0
 DB: 12 Gaps: 0

US-10-659-869A-36 (1-206) x AD003218 (1-968)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
 Db 72 ATGGATAAAAACAACTGTCACACGCTCTCAAGTCTGAAGTGAGAAAAGGCCCTTGG 131
 Qy 21 ThrMetGluGluAspLeuLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTyr 40
 Db 132 ACAATGGAAGAAGACTTGTATCTTCACTCAATATATTCGAAATCATGCGGAGGTGTTTGG 191
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
 Db 192 AACTCTTTGGCCAAAGCTGCTGGTCTCAAGCTACCGGAAGAGTTCCCGGCTAAGGTG 251
 Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 252 CTAACTACCTCCCTGCTGATGTAGAGAGGGATATATACCCGAGGAAACACTTTTG 311
 Qy 81 IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro 100
 Db 312 ATCATGGAGCTTCACGCAAGTGGGGAACAGGTGGTCCAAAATTGCCAAGCATCTACCT 371
 Qy 101 GlyArgThrAspAsnGluLeuLysAsnTyrTyrArgThrArgIleGlnLysHisIleLys 120
 Db 372 GGAAGGACTGATAATAGATCAAGAACTATTGGAGGCAAGGATCCAGAGCATCAAG 431
 Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db 432 CAAGCTGAGAACTTCAGCAACACAGTAGTAAATATCTGAGATAAATGATCCCAAGCT 491
 Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
 Db 492 TGCACTAGCATGTGTCCACCATGCGCAGCCCATGGAGACCTATTCTCCACCCCTGTAT 551
 Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
 Db 552 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTCAATTAATCTCTGATCAATCCAGT 611
 Qy 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTyrSerMetGluAspSerTyrSerMet 200
 Db 612 TGTGTACCAATGACACACACAACTTAATCTATTGGAGCATGGAGGATAGCTGGTCAATG 671
 Qy 201 GlnLeuLeuAsnGlyAsp 206
 Db 672 CAATTACTGAACGGTGAT 689

RESULT 5

ACD07383

ID ACD07383 standard; cDNA; 988 BP.

XX ACD07383;
 AC 07-AUG-2003 (first entry)
 DT Soybean Myb-related transcription factor #2 cDNA.
 DE
 XX Soyben; ss; Myb-related transcription factor; crop improvement;
 KW starch composition; protein composition; plant; gene; grain quality;
 KW disease resistance; pest resistance; oil composition.
 XX
 OS Glycine max.
 XX
 PN US2003024007-A1.
 XX
 PD 30-JAN-2003.
 XX
 XX 14-DEC-2001; 2001US-00021811.
 XX
 PR 02-DEC-1998; 98US-0110609P.
 PR 01-DEC-1999; 99US-00452244.
 XX
 PA (CAHO/) CAHOON R E.
 PA (ODEL/) ODELL J T.
 XX
 PI Cahoon RE, Odell JT;
 DR WPI; 2003-456317/43.
 DR P-PSDB; ABO01756.
 XX
 XX New isolated polynucleotide encoding a Myb-related transcription factor,
 PT useful in plant molecular biology, in particular for the improvement of
 PT crop plants for a variety of traits, including disease and pest
 PT resistance.
 XX
 PS Claim 2; Page 32-33; 61pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a Myb-
 CC related transcription factor polypeptide. The methods and compositions of
 CC the present invention are useful in plant molecular biology, in
 CC particular for the improvement of crop plants for a variety of traits,
 CC including disease and pest resistance and grain quality improvements such
 CC as oil, starch or protein composition. The present sequence represents
 CC cDNA encoding a soybean Myb-related transcription factor
 XX
 SQ Sequence 988 BP; 339 A; 195 C; 182 G; 272 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.16e-107 Length: 988
 Score: 1062.00 Matches: 197
 Percent Similarity: 95.67% Conservative: 2
 Best Local Similarity: 94.71% Mismatches: 7
 Query Match: 93.32% Indels: 2
 DB: 8 Gaps: 2

US-10-659-869A-36 (1-206) x ACD07383 (1-988)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
 Db 61 ATGGATAAAAACAACTGTCACACGCTCTCAAGATCTCTCAAGTGAAGAAAGGACCTTGG 120
 Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTyr 40
 Db 121 ACGATGGAAGAAGACTTGTATCTTGTATCACTATATTGCAAAATCATGGGAAGGTGTTGG 180
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
 Db 181 AATTCTTTGGCCAAAGCTGCTGGTCTCAACAGTACCGGAAGAGTTCCCGCTAAGGTGG 240
 Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 241 CTAACTACCTCCGCTCTGATGTTAGAGAGGGAATATTACCCGAGGAAACACTTTTG 300

QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
 Db |||||
 301 ATCATGGAGCTTCACGCAAGTGGGAAACAGGTGTCCTCAAAATGGCCAGCATCTACCT 360
 QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
 Db |||||
 361 GGTAGGACAGATAATGAGATCAAGAACTATTGGAGGCCAGGATCCAGAAGCACATCAAG 420
 QY 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db |||||
 421 CAAGCTGAGAACTTTTCAGCAACAATTCAGCAATACTCTGAGATAAATGATCACCAGCT 480
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
 Db |||||
 481 AGCACTAGCCATGTTCTTACCATGCTGAACCCATGGAGACCTATTCTCCACCTTTTAT 540
 QY 161 GlnGlyMetLeuGluProPheSerThr---GlnPheProThrIleAsnProAspGlnSer 179
 Db |||||
 541 CAAGGAATGTTAGAGCCATTTTCTTCAATTCAAGTTCCTCCACAAATTAATCTGATCAATCC 600
 QY 180 SerCysCysThrAsnAspAsnAsn---IleAsnTyrTrpSerMetGluAsnSerTrp 198
 Db |||||
 601 AGTTGTTGTACCAATGACAAACACAGCATTAATACTATTGGAGCATGGAGGATATCTGG 660
 QY 199 SerMetGlnLeuLeuAsnGlyAsp 206
 Db |||||
 661 TCAATGCAGTTACTGAACGGGGAT 684

RESULT 6
 ADJ77730
 ID ADJ77730 standard; cdna; 988 BP.
 AC ADJ77730;
 XX
 DT 20-MAY-2004 (first entry)
 DE
 XX
 XX cDNA encoding soybean Myb-related transcription factor #2.
 KW Plant; Myb-related transcription factor; disease resistance; soybean;
 KW gene; ss.
 XX
 OS Glycine max.
 PN
 XX
 XX US2004040057-A1.
 XX
 PD 26-FEB-2004.
 XX
 PF 11-SEP-2003; 2003US-00659869.
 XX
 PR 02-DEC-1998; 98US-0110609P.
 PR 01-DEC-1999; 99US-00452244.
 PR 14-DEC-2001; 2001US-00021811.
 XX
 XX (CAHOON R E.
 PA (FANG/) FANG Y.
 PA (ODELL/) ODELL J T.
 PA (WENG/) WENG Z.
 XX
 PI Cahoon RE, Fang Y, Odell JT, Weng Z;
 XX
 DR WPI; 2004-238121/22.
 DR P-PSDB; ADJ77731.
 XX

Plant Myb transcription factor homologs and their polynucleotide sequences useful in plant genomic research e.g. identifying nucleotide fragments that encode a Myb-related transcription factor polypeptide.

Claim 2; SEQ ID NO 29; 62pp; English.

The present invention relates to the isolation of polynucleotide sequences encoding plant Myb-related transcription factors and the encoding polypeptide sequences. Also disclosed is a chimeric gene encoding all or part of a Myb-related transcription factor linked to

CC suitable regulatory sequences, and a host cell transformed with the
 CC chimeric gene and able to express the chimeric gene resulting in the
 CC production of altered levels of the Myb-related transcription factor in
 CC the transformed cell. The polynucleotide sequences of the invention may
 CC be used to probe a cDNA or genomic library to identify nucleic acid
 CC fragments encoding a Myb-related transcription factor polypeptide and
 CC also for research purposes including generating plants with increased
 CC resistance to disease, etc. The present sequence encodes a plant Myb-
 CC related transcription factor.

XX
 SQ Sequence 988 BP; 339 A; 195 C; 182 G; 272 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.16e-107 Length: 988
 Score: 1062.00 Matches: 197
 Percent Similarity: 95.67% Conservative: 2
 Best Local Similarity: 94.71% Mismatches: 7
 Query Match: 93.32% Indels: 2
 DB: 12 Gaps: 2

US-10-659-869A-36 (1-206) x ADJ77730 (1-988)

QY 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
 Db |||||
 61 ATGGATAAAACAACTGTGCACACCGTCTCAAGATCCTGAAGTGAGAAAAGGACCTTGG 120
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
 Db |||||
 121 ACGATGGAAGAAGACTTGTATCTTGATCAACTATATTGCAATCATGCGGAAGGTGTTGG 180
 QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
 Db |||||
 181 AATCTTTGGCCAAAGCTGCTGCTCAACGTCACCGAAGAGTTCGCGGTAAGGTGG 240
 QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db |||||
 241 CTAACACTACCTCCGTCCTGATGTTAGAGAGGGAATATTACACCCGAGGAACAACCTTTG 300
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
 Db |||||
 301 ATCATGGAGCTTCACGCAAGTGGGAAACAGGTGTCCTCAAAATGGCCAGCATCTACCT 360
 QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
 Db |||||
 361 GGTAGGACAGATAATGAGATCAAGAACTATTGGAGGCCAGGATCCAGAAGCACATCAAG 420
 QY 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db |||||
 421 CAAGCTGAGAACTTTTCAGCAACAATTCAGCAATACTCTGAGATAAATGATCACCAGCT 480
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
 Db |||||
 481 AGCACTAGCCATGTTTTCACCATGGCTGGAACCCATGGAGACCTATTCTCCACCTTTTAT 540
 QY 161 GlnGlyMetLeuGluProPheSerThr---GlnPheProThrIleAsnProAspGlnSer 179
 Db |||||
 541 CAAGGAATGTTAGAGCCATTTTCTTCAATTCAAGTTCCTCCACAAATTAATCTGATCAATCC 600
 QY 180 SerCysCysThrAsnAspAsnAsn---IleAsnTyrTrpSerMetGluAsnSerTrp 198
 Db |||||
 601 AGTTGTTGTACCAATGACAAACACAGCATTAATACTATTGGAGCATGGAGGATATCTGG 660
 QY 199 SerMetGlnLeuLeuAsnGlyAsp 206
 Db |||||
 661 TCAATGCAGTTACTGAACGGGGAT 684

RESULT 7

ADI43073
 ID ADI43073 standard; DNA; 988 BP.
 XX
 AC ADI43073;
 XX
 DT 22-APR-2004 (first entry)

XX	DE	Plant transcription factor polynucleotide #1008.	SQ	Sequence 988 BP; 325 A; 196 C; 193 G; 274 T; 0 U; 0 Other;
XX	KW	transgenic; plant; enhanced tolerance to abiotic stress;	Alignment Scores:	
XX	KW	glyophosphate tolerance; hormone sensitivity; disease resistance;	Pred. No.:	4.38e-88
XX	KW	sugar sensing; flowering; flower structure; stem bifurcation;	Score:	889.50
XX	KW	branching pattern; apical dominance; trichome; stem morphology;	Percent Similarity:	85.65%
XX	KW	root growth; root hair; seed development; cell proliferation;	Best Local Similarity:	81.82%
XX	KW	cell differentiation; premature senescence; necrosis; plant size;	Query Match:	78.16%
XX	KW	leaf morphology; seed morphology; seed biochemistry; root anthocyanin;	DB:	12
XX	KW	plant anthocyanin; light response; shade avoidance; bioinformatic;		
XX	KW	transcription factor; gene; ds.		
XX	OS	Glycine max.	US-10-659-869A-36 (1-206) x ADI43073 (1-988)	
XX	XX	US2004019927-A1.		
XX	XX	29-JAN-2004.		
XX	XX	25-FEB-2003; 2003US-00374780.		
XX	XX	18-APR-2001; 2001US-00837944.		
XX	XX	(SHER/) SHERMAN B K.		
XX	PA	(RIEC/) RIECHMANN J L.		
XX	PA	(JIAN/) JIANG C.		
XX	PA	(HEAR/) HEARD J E.		
XX	PA	(HAKE/) HAAKE V.		
XX	PA	(CREE/) CREELMAN R A.		
XX	PA	(RATC/) RATCLIFFE O.		
XX	PA	(ADAM/) ADAM L J.		
XX	PA	(REUB/) REUBER T L.		
XX	PA	(KEDD/) KEDDIE J.		
XX	PA	(BROU/) BROUN P E.		
XX	PA	(PILG/) PILGRIM M L.		
XX	PA	(DUBE/) DUBELL A N.		
XX	PA	(PINE/) PINEDA O.		
XX	PA	(YUGG/) YU G.		
XX	PI	Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;		
XX	PI	Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;		
XX	PI	Pilgrim ML, Dubell AN, Pineda O, Yu G;		
XX	XX	WPI; 2004-132245/13.		
XX	PT	New transgenic plant comprising a recombinant polynucleotide of any one		
XX	PT	of more than 500 nucleotide sequences, useful in bioinformatic search		
XX	PT	methods.		
XX	PS	Claim 1; SEQ ID NO 1536; 435pp; English.		
XX	XX	The invention describes a transgenic plant comprising a recombinant		
XX	CC	polynucleotide of any one of more than 500 nucleotide sequences fully		
XX	CC	defined in the specification or its complement. The method of the		
XX	CC	invention can be used to produce a plant having altered traits such as:		
XX	CC	enhanced tolerance to abiotic stress; glyophosphate tolerance; hormone		
XX	CC	sensitivity; disease resistance; sugar sensing; early or late flowering;		
XX	CC	altered flower structure, change in stem bifurcations, altered branching		
XX	CC	pattern, reduced apical dominance, reduced trichome density; lack of		
XX	CC	trichomes; reduced ectopic trichome development; altered trichome		
XX	CC	development; increase in trichome number; altered stem morphology;		
XX	CC	increased root growth; increased root hairs; altered seed development;		
XX	CC	altered cell proliferation or cell differentiation; rapid development;		
XX	CC	premature senescence; increased necrosis; increase in seedling or plant		
XX	CC	size; decreased plant size; leaf morphology; seed morphology; seed		
XX	CC	biochemistry; increase in root anthocyanins; increase in plant		
XX	CC	anthocyanins, or alteration in light response or shade avoidance. The		
XX	CC	transgenic plant, polynucleotides and polypeptides are useful in		
XX	CC	bioinformatic search methods. This sequence represents a plant		
XX	CC	transcription factor, and an orthologue of Arabidopsis thaliana		
XX	CC	transcription factors isolated in the invention, that can be used in the		
XX	CC	creation of a transgenic plant with altered traits.		

Qy	1	MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr	20
Db	63	ATGGACAAAAA---CCATGCAACTCATCTCATGATCTCTGAAGTGAAGGACCATGG	119
Qy	21	ThrMetGluGluAspLeuMetAsnTyrIleAlaAsnHisGlyGluValTyr	40
Db	120	ACCATGGAAGAAGACTTGTATTTGATAAATATATTGCAANTCAGGTGAAGGTGTTGG	179
Qy	41	AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr	60
Db	180	AACCTTTAGCCAAAGCTTCTGGCTTAACAGAACGGGAAGAGTTGTCCGACTCCGTTGG	239
Qy	61	LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu	80
Db	240	CTAAACTTACCTTCCTGCTGATGTTAGAAAGAGGAACAATTACACCCGAGGAACAGCTTTTG	299
Qy	81	IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro	100
Db	300	ATCATAGAACTTCATGCAAAAGTGGGCAATAGTGGTCCAAAATTGCAAGCATCTTCCA	359
Qy	101	GlyArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGlnLysHisIleLys	120
Db	360	GGAGAACTGACATGATGATTAAGAACTTCTGGAGAACAGGATCCAAAGACATTAAG	419
Qy	121	GlnAlaGluAsnPheGlnGlnGlnSerSerAsnAsnSerGluLeuAsnAspHisGlnAla	140
Db	420	CAAGCTGAGACTTTCACAAACAT-----GGTAATTTCAGAGAAATATGATCATCAAGCA	473
Qy	141	-----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProPro	158
Db	474	AGCACTAGTAGTACTAGCAAGGTCCACCATGGCAATCCAAATGAGACTTTCTCTCCACCC	533
Qy	159	CysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGln	178
Db	534	TCATACCAAGAACTTTTGAGCCATTCACACTCAATTCCTCCATCACT---GATCAA	590
Qy	179	SerSerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTyr	198
Db	591	TCAAGTTGTTGTACCACCAACCAACCAACCAACTATTGGAGCATCGAGGATATCTG	650
Qy	199	---SerMetGlnLeuLeuAsnGlyAsp	206
Db	651	TCGTCTATGCAATTACTCAATGGAGAT	677
XX	RESULT 8		
XX	ADO03217	ID ADO03217 standard; cDNA; 988 BP.	
XX	AC	ADO03217;	
XX	DT	01-JUL-2004 (first entry)	
XX	DE	Soybean orthologue of Thalecress transcription factor, cDNA #269.	
XX	KW	Soybean; transcription factor; ss; gene; plant; transgenic;	
XX	KW	abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;	
XX	KW	phosphate limitation; potassium limitation; nitrogen limitation;	
XX	KW	hormone sensitivity; disease resistance; sugar sensing; seed germination;	
XX	KW	flowering; inflorescence architectural change;	
XX	KW	meristem cell differentiation; phyllotaxy; apical dominance;	
XX	KW	trichome development; seed development; premature senescence;	

KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.
 XX Glycine max.
 XX US2004045049-A1.
 XX 04-MAR-2004.
 XX 10-APR-2003; 2003US-00412699.
 XX 13-SEP-1999; 99US-00394519.
 PR 21-JAN-2000; 2000US-00489376.
 PR 17-FEB-2000; 2000US-00506720.
 PR 22-MAR-2000; 2000US-00532591.
 PR 22-MAR-2000; 2000US-00533029.
 PR 22-MAR-2000; 2000US-00533030.
 PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-00533648.
 PR 06-APR-2000; 2000WO-US009448.
 PR 16-NOV-2000; 2000US-00713994.
 PR 27-MAR-2001; 2001US-00819142.
 PR 17-APR-2001; 2001US-00837444.
 PR 30-JAN-2002; 2002US-00958131.
 PR 14-JUN-2002; 2002US-00171468.
 PR 09-AUG-2002; 2002US-00225066.
 PR 09-AUG-2002; 2002US-00225067.
 PR 09-AUG-2002; 2002US-00225068.
 PR 17-DEC-2002; 2002US-0434168P.
 PR 25-FEB-2003; 2003US-00374780.
 XX (ZHAN/) ZHANG J.
 PA (FROM/) FROM M E.
 PA (HEAR/) HEARD J E.
 PA (RIEC/) RIECHMANN J L.
 PA (ADAM/) ADAM L J.
 PA (BROU/) BROUN P E.
 PA (PINE/) PINEDA O.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J S.
 PA (YUGG/) YU G.
 PA (JIAN/) JIANG C.
 PA (SAMA/) SAMAHA R S.
 PA (PILG/) PILGRIM M L.
 PA (CREE/) CREELMAN R A.
 PA (DUBE/) DUBELL A N.
 PA (RATC/) RATCLIFFE O.
 PA (KUMI/) KUMIMOTO R.
 PA (SHER/) SHERMAN B K.
 XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
 PI Pinada O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
 PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
 PI Sherman BK;
 XX WPI; 2004-225755/21.
 XX New transgenic plant, useful in developing phenotypes with altered or
 PT improved characteristics or traits.
 XX Claim 1; SEQ ID NO 1631; 213pp; English.
 XX The invention relates to a transgenic plant comprises a recombinant
 CC polynucleotide having a polynucleotide sequence or its complementary
 CC sequence comprising a sequence encoding a polypeptide, that initiates
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
 CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
 CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
 CC plant to grow a progeny plant, an expression cassette (comprising a
 CC constitutive, inducible or tissue-specific promoter and a recombinant
 CC polynucleotide described above), a host cell comprising the expression
 CC cassette, producing a modified plant having a modified trait, identifying
 CC a factor that is modulated by or interacts with a polypeptide encoded by

CC the polynucleotide sequence and identifying at least one downstream
 CC polynucleotide sequence that is subject to a regulatory effect of any of
 CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed development, altered cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed oil
 CC biochemistry, an increase in seed oil content, decrease in seed fatty acid
 CC content, increase in seed fatty acid content, decrease in seed protein
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenyl lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence encodes an
 CC orthologue of a thalecress transcription factor isolated from Soybean.
 XX
 SQ Sequence 988 BP; 325 A; 196 C; 193 G; 274 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.38e-88 Length: 988
 Score: 889.50 Matches: 171
 Percent Similarity: 85.65% Conservative: 8
 Best Local Similarity: 81.82% Mismatches: 23
 Query Match: 78.16% Indels: 7
 DB: 12 Gaps: 5
 US-10-659-869A-36 (1-206) x ADO03217 (1-988)
 QY 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
 DB 63 ATGGACAAAAA---CCATGCAACTCATCTCATCTGAAGTAGAAGAGGACCATGG 119
 QY 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
 DB 120 ACCATGGAGAGACTTGAATCTTGAATAACTATATTGCAATCATCGGTGAAGGTGTGG 179
 QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
 DB 180 AACTCTTTAGCAAAAGCTTCTGGTCTTAAACGAACGGGAAAGAGTTGTCTGACTCCGTTGG 239
 QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 DB 240 CTAACATCACTTCCTCTGATGTTAGAGAGGAACATATTACCCCGAGGAACAGCTTTTG 299
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
 DB 300 ATCATAGACTTCATGCAGAGTGGGGCAATAGTGGTCTCAAAATTCGAAACGACTTCCA 359
 QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120

Db 360 GGAAGAACTGCAATGAGATTAAAGAACTTCGGAGAACCAAGGATCCAAAACACATTAAAG 419
Qy 121 GlnAlaGluAsnPhenGlnGlnSerSerAsnAsnSerGluLeuAsnAsnHisGlnAla 140
Db 420 CAAGCTGAGACTTCACACACACAT-----GGTAATTCAGAGAAATATGATCATCAAGCA 473
Qy 141 -----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerPro 158
Db 474 AGCACTAGTACTAGCAAAAGTTCACCATGTCACATCCAAATGAGACTTTCTCTCCACCC 533
Qy 159 CysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGln 178
Db 534 TCATACCAAGGAACCTTTGAGCCATTCCAACTCAATTCCTTACATCACT---GATCAA 590
Qy 179 SerSerCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp 198
Db 591 TCAAGTTGTTGTACCAACCAACGACACAACTATTGAGCATCGAGGATATCTGG 650
Qy 199 ---SerMetGlnLeuLeuAsnGlyAsp 206
Db 651 TCGTCTATGCAATTACTCAATGGAGAT 677

RESULT 9

ACD07387

ID ACD07387 standard; cDNA; 805 BP.

XX AC ACD07387;

XX DT 07-AUG-2003 (first entry)

XX DE Soybean Myb-related transcription factor #6 cDNA.

XX KW Soybean; ss; Myb-related transcription factor; crop improvement;

XX KW starch composition; protein composition; plant; gene; grain quality;

XX KW disease resistance; pest resistance; oil composition.

XX OS Glycine max.

XX FN US2003024007-A1.

XX PD 30-JAN-2003.

XX PF 14-DEC-2001; 2001US-00021811.

XX PR 02-DEC-1998; 98US-0110609P.

XX PR 01-DEC-1999; 99US-00452244.

XX XX (CAHOON) CAHOON R E.

XX XX (ODEL/) ODELL J T.

XX XX Cahoon RE, Odell JT;

XX XX WPI; 2003-456317/43.

XX XX P-PSDB; ABO01760.

XX PT New isolated polynucleotide encoding a Myb-related transcription factor,

XX PT useful in plant molecular biology, in particular for the improvement of

XX PT crop plants for a variety of traits, including disease and pest

XX PT resistance.

XX PS Claim 2; Page 37-38; 61pp; English.

XX CC The invention relates to an isolated polynucleotide encoding a Myb-

XX CC related transcription factor polypeptide. The methods and compositions of

XX CC the present invention are useful in plant molecular biology, in

XX CC particular for the improvement of crop plants for a variety of traits,

XX CC including disease and pest resistance and grain quality improvements such

XX CC as oil, starch or protein composition. The present sequence represents

XX CC cDNA encoding a soybean Myb-related transcription factor

XX XX Sequence 805 BP; 285 A; 157 C; 155 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.49e-86 Length: 805
Score: 874.50 Matches: 168
Percent Similarity: 85.51% Conservative: 9
Best Local Similarity: 81.16% Mismatches: 23
Query Match: 76.85% Indels: 7
DB: 8 Gaps: 5

US-10-659-869A-36 (1-206) x ACD07387 (1-805)

Qy 4 LysGlnGlnCys---LysThrSerGlnAspProGluValArgLysGlyProTyrThrMet 22
Db 1 AAAAAACCATGCAACTCATCTCATGATCCTGAAGTGAGAAAGGACCATGACCATG 60
Qy 23 GluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSer 42
Db 61 GAAGAAGACTTGATCTTTGATAAACTATATTGCAAAATCACGGTGAAGGTGTTGGAATCC 120
Qy 43 LeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsn 62
Db 121 TTAGCCAAAGCTTCTGGTCTCAAAACGAAACGGAAGAGTTGTGACTCCGTGCTAAAC 180
Qy 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGlnLeuLeuMet 82
Db 181 TACCTTCGTCCTGATGTTAGAAGAGAAACATTACACCCGAGGAACAGCTTTGATCAT 240
Qy 83 GluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
Db 241 GAACCTTCATGCAAAAGTGGGCAATAGGTGTCCTCAAAATTCGAAAGCATCTTCCAGGA 300
Qy 103 ThrAspAsnGluLeuLysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAla 122
Db 301 ACTGACAATGAGATTAGAACTTCTGGAGAACCAAGGATCCAAAAGCACATTAAAGCA 360
Qy 123 GluAsnPhenGlnGlnSerSerAsnAsnSerGluLeuAsnAspHisGlnAla----- 140
Db 361 GAGACTTCACACAAACAT-----GGTAATTCAGAGAAATATGATCATCAAGCAAGCA 414
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
Db 415 AGTACTAGCAAAAGTGTCCACCATGGCACATCCAAATGAGACTTTCTCTCCACCCCTCAT 474
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 180
Db 475 CAAGGAACCTTTGAGCCATTCCAACCTCAATTCCTTACAATCACT---GATCAATCAAG 531
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp---Ser 199
Db 532 TGTGTACCAACCAACCAACCAACCAACTATTGGAGCATCGAGGATATCTGCTGCTCT 591
Qy 200 MetGlnLeuLeuAsnGlyAsp 206
Db 592 ATGCAATTACTCAATGGAGAT 612

RESULT 10

ADJ77738

ID ADJ77738 standard; cDNA; 805 BP.

XX AC ADJ77738;

XX DT 20-MAY-2004 (first entry)

XX XX cDNA encoding soybean Myb-related transcription factor #6.

XX KW Plant; Myb-related transcription factor; disease resistance; soybean;

XX KW gene; ss.

XX OS Glycine max.

XX XX US2004040057-A1.

XX XX 26-FEB-2004.

XX PD

PF 11-SEP-2003; 2003US-00659869.
 XX
 PR 02-DEC-1998; 98US-0110609P.
 PR 01-DEC-1999; 99US-00452244.
 PR 14-DEC-2001; 2001US-00021811.
 XX
 PA (CAHO/) CAHOON R E.
 PA (FANG/) FANG Y.
 PA (ODEL/) ODELL J T.
 PA (WENG/) WENG Z.
 XX
 PI Cahoon RE, Fang Y, Odell JT, Weng Z;
 XX WPI; 2004-238121/22.
 DR P-PSDB; ADJ77739.
 XX
 PT Plant Myb transcription factor homologs and their polynucleotide
 PT sequences useful in plant genomic research e.g. identifying nucleotide
 PT fragments that encode a Myb-related transcription factor polypeptide.
 XX
 PS Claim 2; SEQ ID NO 37; 62pp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding plant Myb-related transcription factors and the
 CC encoding polypeptide sequences. Also disclosed is a chimeric gene
 CC encoding all or part of a Myb-related transcription factor linked to
 CC suitable regulatory sequences, and a host cell transformed with the
 CC chimeric gene and able to express the chimeric gene resulting in the
 CC production of altered levels of the Myb-related transcription factor in
 CC the transformed cell. The polynucleotide sequences of the invention may
 CC be used to probe a cDNA or genomic library to identify nucleic acid
 CC fragments encoding a Myb-related transcription factor polypeptide and
 CC also for research purposes including generating plants with increased
 CC resistance to disease, etc. The present sequence encodes a plant Myb-
 CC related transcription factor.
 XX
 SQ Sequence 805 BP; 285 A; 157 C; 155 G; 208 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,498-86 Length: 805
 Score: 874.50 Matches: 168
 Percent Similarity: 85.51% Conservative: 9
 Best Local Similarity: 81.16% Mismatches: 23
 Query Match: 76.85% Indels: 7
 DB: 12 Gaps: 5
 US-10-659-869A-36 (1-206) x ADJ77738 (1-805)
 QY 4 LysGlnGlnCys---LysThrSerGlnAspProGluValArgLysGlyProTrpThrMet 22
 DB 1 AAAAAACCATGCAACTCATCTCATCTGATCTCTGAAGTGAGAAAGGACCATG 60
 QY 23 GluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSer 42
 DB 61 GAAGAAGACTTGATCTTGATAAACTATATTGCAAAATCAGCGTGAAGGTGTTGGAACTCC 120
 QY 43 LeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsn 62
 DB 121 TTAGCCAAAGCTTCGTGCTCTCAACGACGAGAAAGAGTTCGACTCCGTGGCTAAAC 180
 QY 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeuMet 82
 DB 181 TACCTTCGTCCTGATGTTAGAGAGGAAACATTACACCGAGGAACAGCTTTTGATCATA 240
 QY 83 GluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
 DB 241 GAACTTCATGCAAGTGGGGCAATAGGTGTCAAAATTCGAAAGCATCTTCAGGAAGA 300
 QY 103 ThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAla 122
 DB 301 ACTGACATGAGATTAAAGAACTTCGTGGAGAACAGATCCAAAGACACATTAAGCAAGCT 360
 QY 123 GluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla----- 140

DB 361 GAGACTTCACAAACAT-----GGTAATTCAGAGAATAATGATCATCAAGCAAGCACT 414
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
 DB 415 AGTACTAGCAAAAGTGTCCACCATGCGCATCCAAATGAGACTTTTCTCTCCACCTCATAC 474
 QY 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
 DB 475 CAAGGAACCTTTTGAGCCATTCACACCTCAATTCCTACATCACT---GATCAATCAAGT 531
 QY 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp---Ser 199
 DB 532 TGTGTATACCAACCAACGACACCAACAACAACTATTGGAGCATCGAGGATATCTGGTGGTCT 591
 QY 200 MetGlnLeuLeuAsnGlyAsp 206
 DB 592 ATGCAATTACTCAATGGAGAT 612
 RESULT 11
 ACD07385
 ID ACD07385 standard; cDNA; 910 BP.
 XX
 AC ACD07385;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Soybean Myb-related transcription factor #4 cDNA.
 XX
 KW Soyben; ss; Myb-related transcription factor; crop improvement;
 KW starch composition; protein composition; plant; gene; grain quality;
 KW disease resistance; pest resistance; oil composition.
 XX
 OS Glycine max.
 XX
 FN US2003024007-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 14-DEC-2001; 2001US-00021811.
 XX
 PR 02-DEC-1998; 98US-0110609P.
 PR 01-DEC-1999; 99US-00452244.
 XX
 PA (CAHO/) CAHOON R E.
 PA (ODEL/) ODELL J T.
 XX
 PI Cahoon RE, Odell JT;
 XX
 DR WPI; 2003-456317/43.
 DR P-PSDB; ABO01758.
 XX
 PT New isolated polynucleotide encoding a Myb-related transcription factor,
 PT useful in plant molecular biology, in particular for the improvement of
 PT crop plants for a variety of traits, including disease and pest
 PT resistance.
 XX
 PS Claim 2; Page 35; 61pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a Myb-
 CC related transcription factor polypeptide. The methods and compositions of
 CC the present invention are useful in plant molecular biology, in
 CC particular for the improvement of crop plants for a variety of traits,
 CC including disease and pest resistance and grain quality improvements such
 CC as oil, starch or protein composition. The present sequence represents
 CC cDNA encoding a soybean Myb-related transcription factor
 XX
 SQ Sequence 910 BP; 326 A; 164 C; 170 G; 246 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 1,618-81 Length: 910
 Score: 829.50 Matches: 168
 Percent Similarity: 83.89% Conservative: 9

Best Local Similarity: 79.62% Mismatches: 23
Query Match: 72.89% Indels: 12
DB: 8 Gaps: 7

US-10-659-869A-36 (1-206) x ACD07385 (1-910)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTirp 20
Db 37 ATGGACAAAAA---CCATCGACTCATCTCATGATCCAGAGTGAGAGGACCATGG 93
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTirp 40
Db 94 ATCATGAGAGAGACTTGTATTAACATATATTCGAATACAGGTGAGGTGGTGG 153
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTirp 60
Db 154 AATTCCTTTAGCAAAAGCTTCTGGTCTTAACCAAGGAGGAGTGTGCACTCCGTTGG 213
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 214 CTAAACTACCTTCGTCCTGATGTTAGAGAGAGAAACATTCACCCGAGAACACAGCTTTTG 273
Qy 81 IleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisLeuPro 100
Db 274 ATCATAGACTTCATGCAAGTGGGCAATAGGTGGTCCAAANTTGCAGACATCTTCCA 333
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTirpArgThrArgIleGlnLysHisIleLys 120
Db 334 GGAAGAACTGACATGAGATTAAAGACTTCTGGAGAACTAGGATCCAGAGACACATTAA 393
Qy 121 GlnAlaGluAsnPheGlnGln-----GlnSerSerAsnAsnSerGluIleAsnAspHis 138
Db 394 CAAGCTGAGACTTCACAAACACATGGTAATTCATCAGAGATAGTAAT---AATGATCAT 450
Qy 139 GlnAla-----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSer 156
Db 451 CAACCAAGCAATAGACATAGCAAGGTGTCACCACTGACATCCCAATGAGACTTCTCT 510
Qy 157 ProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnPro 176
Db 511 TCACCTCATACCAAGCAACTTTTGGCCATTTCACCTCAATTCCTACATCAAT---566
Qy 177 AspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTirpSerMetGluAsp 196
Db 567 GATCAATCAAGTTCTGTACCAGCAACAAC-----AACTATTGAGCATCGAGGAT 617
Qy 197 SerTirp---SerMetGlnLeuLeuAsnGlyAsp 206
Db 618 ATCTGGTGTCTATGCAATTAATCTCAATGGAGAT 650

RESULT 12
ADJ77734
ID ADJ77734 standard; cDNA; 910 BP.
XX AC ADJ77734;
XX DT 20-MAY-2004 (first entry)
XX DE cDNA encoding soybean Myb-related transcription factor #4.
KW Plant; Myb-related transcription factor; disease resistance; soybean;
KW gene; ss.
XX OS Glycine max.
XX PN US2004040057-A1.
XX PD 26-FEB-2004.
XX PF 11-SEP-2003; 2003US-00659869.
XX PR 02-DEC-1998; 98US-0110609P.
PR 01-DEC-1999; 99US-00452244.

PR 14-DEC-2001; 2001US-00021811.
XX (CAHO/) CAHOON R E.
PA (FANG/) FANG Y.
PA (ODEL/) ODELL J T.
XX (WENG/) WENG Z.
XX Cahoon RE, Fang Y, Odell JT, Weng Z;
DR WPI; 2004-238121/22.
DR P-PSDB; ADJ77735.
XX Plant Myb transcription factor homologs and their polynucleotide
PT sequences useful in plant genomic research e.g. identifying nucleotide
PT fragments that encode a Myb-related transcription factor polypeptide.
XX Claim 2; SEQ ID NO 33; 62pp; English.
XX The present invention relates to the isolation of polynucleotide
CC sequences encoding plant Myb-related transcription factors and the
CC encoding polypeptide sequences. Also disclosed is a chimeric gene
CC encoding all or part of a Myb-related transcription factor linked to
CC suitable regulatory sequences, and a host cell transformed with the
CC chimeric gene and able to express the chimeric gene resulting in the
CC production of altered levels of the Myb-related transcription factor in
CC the transformed cell. The polynucleotide sequences of the invention may
CC be used to probe a cDNA or genomic library to identify nucleic acid
CC fragments encoding a Myb-related transcription factor polypeptide and
CC also for research purposes including generating plants with increased
CC resistance to disease, etc. The present sequence encodes a plant Myb-
CC related transcription factor.
XX SQ Sequence 910 BP; 326 A; 164 C; 170 G; 246 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 1,61e-81 Length: 910
Score: 829.50 Matches: 168
Percent Similarity: 83.89% Conservative: 9
Best Local Similarity: 79.62% Mismatches: 23
Query Match: 72.89% Indels: 12
DB: Gaps: 7
US-10-659-869A-36 (1-206) x ADJ77734 (1-910)
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTirp 20
Db 37 ATGGACAAAAA---CCATCGACTCATCTCATGATCCAGAGTGAGAGGACCATGG 93
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTirp 40
Db 94 ATCATGAGAGAGACTTGTATTAACATATATTCGAATACAGGTGAGGTGGTGG 153
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTirp 60
Db 154 AATTCCTTTAGCAAAAGCTTCTGGTCTTAACCAAGGAGGAGTGTGCACTCCGTTGG 213
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 214 CTAAACTACCTTCGTCCTGATGTTAGAGAGAGAAACATTCACCCGAGAACACAGCTTTTG 273
Qy 81 IleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisLeuPro 100
Db 274 ATCATAGACTTCATGCAAGTGGGCAATAGGTGGTCCAAANTTGCAGACATCTTCCA 333
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTirpArgThrArgIleGlnLysHisIleLys 120
Db 334 GGAAGAACTGACATGAGATTAAAGACTTCTGGAGAACTAGGATCCAGAGACACATTAA 393
Qy 121 GlnAlaGluAsnPheGlnGln-----GlnSerSerAsnAsnSerGluIleAsnAspHis 138
Db 394 CAAGCTGAGACTTCACAAACACATGGTAATTCATCAGAGATAGTAAT---AATGATCAT 450
Qy 139 GlnAla-----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSer 156

```
Db 451 CAAGCAAGCAATAGCACTAGCAAGGTGCCACATGGCACATCCAAATGAGACTTTCTCT 510
Qy 157 ProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnPro 176
Db 511 TCACCTCATACCAAGCACTTTTGAGGCAATTTCA-ACCTCAATCTCTACATCAAT--- 566
Qy 177 AspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAsp 196
Db 567 GATCAATCAAGTTGTTGTACCAAGCAACAAC-----NACTATTGGAGCATCGAGGAT 617
Qy 197 SerTrp---SerMetGlnLeuLeuAsnGlyAsp 206
Db 618 ATCTGGTCGTCTATCAATTAATCAATGGAGAT 650

RESULT 13
ID ADI43075
XX ADI43075 standard; DNA; 969 BP.
XX AC ADI43075;
XX DT 22-APR-2004 (first entry)
XX DE Plant transcription factor polynucleotide #1010.
XX KW transgenic; plant; enhanced tolerance to abiotic stress;
KW phosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; gene; ds.
XX Glycine max.
XX US2004019927-A1.
XX PN 29-JAN-2004.
XX PD 25-FEB-2003; 2003US-00374780.
XX PF 18-APR-2001; 2001US-00837944.
XX PR (SHER/) SHERMAN B K.
PA (RIE/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAR/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
XX
XX New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX
XX Claim 1; SEQ ID NO 1538; 435pp; English.
XX
```

```
CC The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produced a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanins, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant in
CC transcription factor, and an orthologue of Arabidopsis thaliana
CC creation of a transgenic plant with altered traits.
XX
SQ Sequence 969 BP; 311 A; 188 C; 200 G; 270 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.: 2,49e-79 Length: 969
Score: 810.00 Matches: 158
Percent Similarity: 86.60% Conservative: 10
Best Local Similarity: 81.44% Mismatches: 24
Query Match: 71.18% Indels: 3
DB: 12 Gaps: 2
```

US-10-659-869A-36 (1-206) x ADI43075 (1-969)

```
Qy 12 AspProGluValArgLysGlyProThrMetGluGluAspLeuIleLeuMetAsnTyr 31
Db 26 GATCCTGCCCTGTATTTCGGGCCAGG-ACAATGAAGAAGACTTAATCTTGATCACCTAT 84
Qy 32 IleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaGlyLeuLysArg 51
Db 85 ATTGCCAATCACGGGGAAGGGGTTTGGAACTCTTTGGCAAGGCTGTGGGACTTTAAACGT 144
Qy 52 AsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgGly 71
Db 145 ACCGGAAGAAGATTCCGGCTTCGGTAGCTAACTACCTACCTCTCTGTATGATGAGAGGG 204
Qy 72 AsnIleThrProGluGluGlnLeuLeuMetGluLeuHisAlaLysTrpGlyAsnArg 91
Db 205 AATATTACCCCGAGGAAACAGCTTTTGATCATGGAACCTTCATGCAAGTGAGGACACAGG 264
Qy 92 TrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrp 111
Db 265 TGGTCCAAAATTGCCAAGCATCTACCCGGAAGGACTGATATATGAGATTAAAGACTACTGG 324
Qy 112 ArgThrArgIleGlnLysHisIleLysGlnAla---GluAsnPheGlnGlnSerSer 130
Db 325 AGGACAAAGATCCAGAAGCACCTCAAGCAAGCTTCCAGCAGCTTCCAGCAACAGAGTAGT 384
Qy 131 AsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGlu 150
Db 385 AATCTGAGATAATTTATCATCCCAAGCTTGCACTAGCCAAAGTGTCCCATGGCGAG 444
Qy 151 ProMetGluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGln 170
Db 445 CCCATAGAACTATTCTCCACCCAGTTATCAAGGAATGTTAGATCCATTTCAATTCCAG 504
Qy 171 PheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsn 190
Db 505 TTCCCAACA---AATCTCTCATCTTCTAGTTGTTGTACCAATGACGACGACCAACAAC 561
Qy 191 TyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsn 204
Db 562 TATTGGAGCATGGAGGATATCTGTCAATGCAATTAGCCAAT 603
```



```
QY 32 lleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArg 51
|
|
|
Db 85 ATTGCATACAGGGGAGGGGTTGGAACTCTTTGGCCAGAGGCTGCTGGACTTAAAGGT 144
|
|
|
QY 52 AsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArgGly 71
|
|
|
Db 145 ACCGGAAAGAGTTCGCGGCTTCGGTAGCTAAACTACCTACCTACCTGCTGATGTTAGAGAGGG 204
|
|
|
QY 72 AsnIleThrProGluGluGlnLeuLeuLeuMetGluLeuHisAlaLysTrpGlyAsnArg 91
|
|
|
Db 205 AATATATACACCGAGGACAGCTTTTGATCATGGAACTTCATGCAAGTGGAGCAACAGG 264
|
|
|
QY 92 TrpSerLysIleAlaLysHisLeuProGlyArgTrpAspAsnGluIleLysAsnTyrTrp 111
|
|
|
Db 265 TGTCTCAAAATTCACAGCATCTACCCGAGGAGCTGATATAGATTTAAGAACTACTGG 324
|
|
|
QY 112 ArgThrArgIleGlnLysHisIleLysGlnAla---GluAsnPheGlnGlnSerSer 130
|
|
|
Db 325 AGGCAAGGATCCAGAAGCACCTCAAGCAAGCTTCAGCAGCTTCAGCAACAGAGTAGT 384
|
|
|
QY 131 AsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGlu 150
|
|
|
Db 385 AATTCGTGAGATAATTTATCATCCCAAGCTTGCACTAGCCCAAGTGTCCACCATGGCGCAG 444
|
|
|
QY 151 ProMetGluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGln 170
|
|
|
Db 445 CCATAGAAACCTATCTCCACCAGTTATCAAGGAATGTAGATCCATTTCAATTCCAG 504
|
|
|
QY 171 PheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsn 190
|
|
|
Db 505 TTCCCAACA---AATCCTCATCTAGTTGTTGTGTACCAATGACGACGACCAACAAC 561
|
|
|
QY 191 TyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsn 204
|
|
|
Db 562 TATTGGAGCATGGAGATATCTGGTCAATGCCAATTAGCCAAT 603
|
|
|
RESULT 15
ADD17714
ID ADD17714 standard; DNA; 681 BP.
XX
AC ADD17714;
XX
XX 15-JAN-2004 (first entry)
XX
DE DNA (SeqID 1782) that confers an altered visual phenotype in plants.
XX
KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;
KW bleaching; etching; wet leaf; stunting; elongation; texture;
KW agronomic trait; growth regulation; dwarf variety; insect resistance;
KW heat stress; transgenic.
XX
XX Unidentified.
XX
OS WO2003020741-A1.
XX
PN 13-MAR-2003.
XX
PD 30-AUG-2002; 2002WO-US027880.
XX
PF 31-AUG-2001; 2001US-0316326P.
XX
PR (DOWC ) DOW CHEM CO.
XX
PA (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Croasley R, Skokut T, Ruegger M, Larrinua I, Shukla V;
XX WPI; 2003-300858/29.
XX
DR Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
XX sativa, Saccharomyces cerevisiae, and Papaver thoeas, useful for
PT conferring altered visual phenotypes in plants.
XX
```

PS Claim 1; SEQ ID NO 1782; 517bp; English.

XX This invention relates to the identification and isolation of novel
CC nucleic acid molecules that confer altered visual phenotypes in plants.
CC Specifically, it refers to modifications of plant architecture and/or
CC leaf surface features in plants, such as chlorotic, bleaching, etching,
CC wet leaf, stunting, elongation and texture phenotypes, which are thought
CC will be agronomic traits beneficial to the farmer. As such, these novel
CC phenotypes can affect growth regulation i.e. useful for creating dwarf
CC varieties, exhibit resistance to insects or heat stress, confer changes
CC in pigment content such that plants have enhanced vitamin production or
CC delayed senescence and also for example produce plants that control the
CC production of ethylene. Furthermore, the present invention comprises
CC generating transgenic plants, as well as reproducibly altering the visual
CC phenotype of plant seeds, plant tissues and plant cells containing the
CC polynucleotides described herein. This polynucleotide is a homologue of a
CC DNA sequence that confers an altered visual phenotype when expressed in
XX plants, the method of the invention.

XX Sequence 681 BP; 221 A; 130 C; 137 G; 193 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,98e-74 Length: 681
Score: 760.00 Matches: 151
Percent Similarity: 79.31% Conservative: 10
Best Local Similarity: 74.38% Mismatches: 18
Query Match: 66.78% Indels: 24
DB: 10 Gaps: 6

US-10-659-869A-36 (1-206) x ADD17714 (1-681)

```
QY 10 SerGlnAspProGluValArgLysGlyProTrpThrMetGluGluAspLeuIleLeuMet 29
|
|
|
Db 91 TCTCAAGATGTTGAAGTGAGGAAAGGACCTTGGACTATGGAGAGGATTTAATTCATC 150
|
|
|
QY 30 AsnTyrIleAlaAsnHisGlyGlyValTrpAsnSerLeuAlaLysAlaGlyLeu 49
|
|
|
Db 151 AACTACATGCTAATCATGCTGAAGGTGTTTGGAAATTCCTTAGCAAAATCTGCTGCTC 210
|
|
|
QY 50 LysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArg 69
|
|
|
Db 211 AACGAAACCGGAAAGAGCTGTCGACTCCGGTGGCTTAAATATCTCCGCCCTGATGTCGG 270
|
|
|
QY 70 ArgGlyAsnIleThrProGluGluGlnLeuIleMetGluLeuHisAlaLysTrpGly 89
|
|
|
Db 271 AGGGGAATATTACACCTGAAGAACAACTTTTGATATATGGAACATGCTAAGTGGGGA 330
|
|
|
QY 90 AsnArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsn 109
|
|
|
Db 331 AACAGGTGCTCAAAATTCGAAAGCATTTGCCAGGAAGAACAGATACAGATATAAAAT 390
|
|
|
QY 110 TyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnSer 129
|
|
|
Db 391 TATTGGAGGACAAGAAATTCAGAACACATAAAGCAAGCAGAAACATGAATGGACACAGA 450
|
|
|
QY 130 SerAsnAsnSerGluIleAsnAspHisGln---AlaSerThrSerHisValSerThrMet 148
|
|
|
Db 451 GCT---AATTCGTGAGCAAAATGATCATCAAGAGGAGCAGTAGCCATATGTCGCT--- 504
|
|
|
QY 149 AlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuLeu--- 165
|
|
|
Db 505 GCTGTCGACGACAGACTTACTCTCCAACTTCTCACTCTGCTAATATATTATGACACAACTTT 564
|
|
|
QY 166 -----ProPheSerThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThr 183
|
|
|
Db 565 CAAGGACCTTTTCTCACATGAA-----ACCAACACACACACACACACACACACAC 588
|
|
|
QY 184 AsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeu 203
|
|
|
Db 589 AATGAC-----AACATTTGGAGCATGGAGGATATCTGGTCCATGCAATTCCTT 636
|
|
|
QY 204 AsnGlyAsp 206
|
|
|
```


Db 637 AACGGCGAT 645

Search completed: September 28, 2005, 16:04:56
Job time : 530 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2005, 13:31:32 ; Search time 178 Seconds
(without alignments)
1893.669 Million cell updates/sec

Title: US-10-659-869A-36
Perfect score: 1138
Sequence: 1 MDKKQCKTSQDPVRKGPW.....NNINYSMEDSWSQLNGD 206

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10659869/runat 26092005 143445 11388/app query.fasta_1.391
-DB=Issued_Patents_NA -QWMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10659869 @CGN 1 1 69 @runat 26092005 143445 11388 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759.5	66.7	862	US-09-640-211A-2076	Sequence 2076, Ap
2	653.5	57.4	1033	US-09-533-029-87	Sequence 87, Appl
3	625.5	55.0	907	US-09-640-211A-2077	Sequence 2077, Ap
4	540	47.5	673	US-09-640-211A-328	Sequence 328, App
5	478	42.0	524	US-09-640-211A-1934	Sequence 1934, Ap
6	475	41.7	1347	US-09-533-029-39	Sequence 39, Appl
7	421	37.0	389	US-09-640-211A-283	Sequence 283, App
8	421	37.0	417	US-09-640-211A-1496	Sequence 1496, Ap
9	410.5	36.1	516	US-09-640-211A-483	Sequence 483, App
10	410.5	36.1	516	US-09-640-211A-588	Sequence 588, App
11	405.5	35.6	542	US-09-640-211A-565	Sequence 565, App
12	404.5	35.5	389	US-09-640-211A-1559	Sequence 1559, Ap

13	404	35.5	1081	4	US-09-453-387A-5	Sequence 5, Appli
14	397.5	34.9	1209	4	US-09-533-029-85	Sequence 85, Appl
15	396.5	34.8	1006	4	US-09-453-387A-1	Sequence 1, Appli
16	396	34.8	2220	3	US-08-997-251-1	Sequence 1, Appli
17	392	34.4	1218	4	US-09-640-211A-2086	Sequence 2086, Ap
18	390	34.3	526	4	US-09-640-211A-1985	Sequence 1985, Ap
19	386	33.9	2352	3	US-08-997-251-3	Sequence 3, Appli
20	383	33.7	573	4	US-09-640-211A-2065	Sequence 2065, Ap
21	382.5	33.6	1344	2	US-08-722-626B-1	Sequence 1, Appli
22	381	33.5	1151	4	US-09-453-387A-3	Sequence 3, Appli
23	378.5	33.3	1046	4	US-09-533-029-61	Sequence 61, Appl
24	378	33.2	648	4	US-09-640-211A-2036	Sequence 2036, Ap
25	378	33.2	1285	4	US-09-640-211A-2085	Sequence 2085, Ap
26	377.5	33.2	462	4	US-09-640-211A-1765	Sequence 1765, Ap
27	377.5	33.2	464	4	US-09-640-211A-1736	Sequence 1736, Ap
28	377.5	33.2	511	4	US-09-640-211A-1648	Sequence 1648, Ap
29	377.5	33.2	551	4	US-09-640-211A-1827	Sequence 1827, Ap
30	377	33.1	420	4	US-09-640-211A-494	Sequence 494, App
31	377	33.1	918	4	US-09-533-029-93	Sequence 93, Appl
32	375	33.0	476	4	US-09-640-211A-1987	Sequence 1987, Ap
33	373.5	32.8	636	4	US-09-640-211A-2075	Sequence 2075, Ap
34	373	32.8	786	4	US-09-640-211A-2077	Sequence 2077, Ap
35	370	32.5	338	4	US-09-640-211A-1823	Sequence 1823, Ap
36	369	32.4	513	4	US-09-640-211A-519	Sequence 519, App
37	368.5	32.4	446	4	US-09-640-211A-1713	Sequence 1713, Ap
38	368	32.3	442	4	US-09-640-211A-1693	Sequence 1693, Ap
39	368	32.3	469	4	US-09-640-211A-492	Sequence 492, App
40	367	32.2	424	4	US-09-640-211A-1965	Sequence 1965, Ap
41	367	32.2	504	4	US-09-640-211A-502	Sequence 502, App
42	366	32.2	944	4	US-09-453-387A-7	Sequence 7, Appli
43	365	32.1	1678	4	US-09-640-211A-2074	Sequence 2074, Ap
44	363.5	31.9	495	4	US-09-640-211A-2032	Sequence 2032, Ap
45	363	31.9	395	4	US-09-640-211A-508	Sequence 508, App

ALIGNMENTS

RESULT 1
US-09-640-211A-2076
; Sequence 2076, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2076
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2076

Alignment Scores: 5.79e-89 Length: 862
Pred. No.: 759.50 Matches: 152
Score: 76.30% Conservative: 9
Percent Similarity: 76.30%
Best Local Similarity: 72.04% Mismatches: 29
Query Match: 66.74% Indels: 21
DB: 4 Gaps: 5

US-10-659-869A-36 (1-206) x US-09-640-211A-2076 (1-862)

Qy 1 MetAspLysLys-----GlnGlnCysLysThrSerGlnAspProGluValArgLysGly 18
Db 31 ATGCACGAAGAACCCAGACGACGACGATGTCCTCCCAAGATGTCGAGGTGAGAAAGGG 90

```
Qy 19 ProTrpThrMetGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
Db 91 CGGTGGACGATGGAAGAGATCTCATCTCACTACATAGCAATACACGGCGAAGGC 150
Qy 39 ValTrpAsnSerLeuAlaIysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db 151 AGTTGGAACTCCCTAGCCAAAAGCTCTGGTCTTAAACACGTACCGGAAGAGTTGTGGCTC 210
Qy 59 ArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGln 78
Db 211 CGGTGGCTGAACATATCTGGACCCGACGTCGGAGAGGCAACATACACTACTGAGGAGCAG 270
Qy 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHis 98
Db 271 CTCCTGATCATGGAACATGTCATGCCAAGTGGGAAACAGGTGCTCTAAATTCGAAAGCAT 330
Qy 99 LeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHis 118
Db 331 CTTCCCGGAAGGACTGACAATGAGATAAAGAACTTCTGGAGGACTAGAAATCCAAAAGCAC 390
Qy 119 IleLysGlnAlaGluAsnPheGlnGlnSerSerAsnSerGluIleAsnAspHis 138
Db 391 ATCAAGCAAGCAGAGGCTTCTCGTCCAGAGCTCCGAGATGAGTGAT----- 438
Qy 139 GlnAlaSerThrSerHisValSerThrMetAlaGluProMetGluMetTyr---SerPro 157
Db 439 CAAGCAAGCAACAAAGCCACATGTCAGATGCCAGAGCCGATGGAGACCTACGACTCACCG 498
Qy 158 ProCysTyrGlnGly-----MetLeuGluProPheSerThrGlnPheProThrIleAsn 175
Db 499 CGGTCAATCCAGGCAACAACAATGAGGCGCTTTGCCGGTGAATTTGTCGGTC----- 552
Qy 176 ProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGlu 195
Db 553 -----GAGTCRAATGAAGCCTACTTGAGAGCATGGAC 582
Qy 196 AspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db 583 GATCTTTGGTCTATGTCAGTTACTCAATGGGAT 615
RESULT 2
US-09-533-029-87
; Sequence 87, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1033
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G221
US-09-533-029-87
```

```
Alignment Scores:
Pred. No.: 5,08e-75 Length: 1033
Score: 653.50 Matches: 137
Percent Similarity: 71.83% Conservative: 16
Best Local Similarity: 64.32% Mismatches: 45
Query Match: 57.43% Indels: 15
DB: 4 Gaps: 6
US-10-659-869A-36 (1-206) x US-09-533-029-87 (1-1033)
Qy 9 ThrSerGlnAspProGluValArgLysGlyProTrpThrMetGluGluAspLeuLeu 28
Db 154 TCATCAGCAGAGCAGAGAGTGAGAAAAGCAGCATGGACGATGGAAGAAGATCTTATCTT 213
Qy 29 MetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaIysAlaGly 48
Db 214 ATCAACTATATCGCCAACCCAGCGGATGGTGTGGAAATCTCTCGCAAAATCTCGAGT 273
Qy 49 LeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspVal 68
Db 274 CTAAACGAAACCGGAAAGTTGCCGCTCCGCTGCTGAATCTCTCGCCCGCACGTA 333
Qy 69 ArgArgGlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrp 88
Db 334 CGACGGGAAACATCACTCCAGAAAGCAACTTATCATCATGGAACCTTCATGCTAAGTGG 393
Qy 89 GlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluLys 108
Db 394 GGAACAGGTGGTGGAAATCGCAAAACATCTTCCAGGAAGACGGAACAGAGATCAA 453
Qy 109 AsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGlu-----AsnPhe 125
Db 454 AATTCTGTAGACACAGAAATTCAAAATACATCAAGCAATCGGATGTAACAACACATCG 513
Qy 126 GlnGlnGlnSerSerAsnSerGluIleAsnAspHisGlnAlaSerThrSerHis--- 144
Db 514 TCCGTTGGATCTCATCATAGTCAAGATCAACGATCAAGCTGCAAGCACGTCGAGCCAT 573
Qy 145 -----ValSerThrMetAlaGluProMetGluMetTyrSer-----ProProCysTyr 160
Db 574 AATGCTTTTGTACACAGATCAAGCATGGAGACTTAATCTCTCTACACCGACATCATAT 633
Qy 161 GlnGlyMetLeuGluProPhe-----SerThrGlnPheProThrIleAsn 175
Db 634 CAACATACCATATGGAATTCAACTATGTAATTCGGCGCGGAGTGCAGCAACC 693
Qy 176 ProAspGlnSerSerCysCysThr---AsnAspAsnAsnIleAsnTyrTrpSerMet 194
Db 694 GTGGATTATCCAGTACCAGTACCGTTCATGATCAACCGGTCGAAACACTATTGGGGCATG 753
Qy 195 GluAspSerTrp---SerMetGlnLeuLeuAsnGlyAsp 206
Db 754 GATGATATTGGTCAATCAATGATTTAATGATGTAAT 792
RESULT 3
US-09-640-211A-2077
; Sequence 2077, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2077
; LENGTH: 907
```

```
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2077

Alignment Scores:
Pred. No.: 1.84e-71 Length: 907
Score: 625.50 Matches: 134
Percent Similarity: 69.68% Conservative: 20
Best Local Similarity: 60.63% Mismatches: 28
Query Match: 54.96% Indels: 39
DB: 4 Gaps: 10

US-10-659-869A-36 (1-206) x US-09-640-211A-2077 (1-907)
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
Db 4 ATGGACAAGAAA---CCGTGCTACCGCAGCAGGATCTCAAGTGAGAAAAGGACCATGG 60
Qy 21 ThrMetGluGluAspLeuLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTyr 40
Db 61 ACTCTTGAAGAGGATCTGATTCGTGAGCTACATAGCAACCATGGCGAAGGTGTATGG 120
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
Db 121 AACTCTCTAGCTAAGCTGTGGTCTTCAACGCAACCGAAGAGCTGCCGGCTGCGATGG 180
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 181 TTGNATTTACCTCCGACCGGATGTCGGAGAGGTATATCACTCCCGAAGAACAGCTTTG 240
Qy 81 IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro 100
Db 241 ATCATACACCTGCAATCCATGTGGGAAACAGGTGGTCCGAAATCGCGAAGCATTTGCCG 300
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGlnLysHis---Ile 119
Db 301 GGGAGGAGCCCAATGAATAAAGAACTACTGGAGGACCAAGATCCAAAAGACATAATC 360
Qy 120 LysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAsp----- 137
Db 361 AAGCAATCGGNA-----ACTGAGATCAACGATCTCACT 393
Qy 138 -----HisGlnAlaSerThrSerHisValSer---ThrMetAlaGluPro 151
Db 394 ATTCCTCCATCATCCGCAACGCATGCACATCATCGCGGGTCTCGCGCAATACA 453
Qy 152 MetGluMet---TyrSerProCysTyrGlnGly-----MetLeuGlu 165
Db 454 ATCGAGATCGCCTCTTCCACCGTCGGATCAAGCGCGCTCCGCGGAGACTATGCTC--- 510
Qy 166 ProPheSerThrGlnPheProThrIleAsnProAspGlnSerSerCysThrAsnAsp 185
Db 511 -----TCGGCCCTTCTCTGCCCAAGAACCGAACACAGCGCTTGC----- 552
Qy 186 AsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMetGln---LeuLeuAsn 204
Db 553 -----TGGAGCGTGGAGGATCTCTGGCCCAATACATCACTCAATTAATAGC 594
Qy 205 Gly 205
Db 595 GGC 597

RESULT 4
US-09-640-211A-328
; Sequence 328, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 328
; LENGTH: 673
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-328

Alignment Scores:
Pred. No.: 1.58e-60 Length: 673
Score: 540.00 Matches: 113
Percent Similarity: 61.99% Conservative: 24
Best Local Similarity: 51.13% Mismatches: 36
Query Match: 47.45% Indels: 48
DB: 4 Gaps: 6

US-10-659-869A-36 (1-206) x US-09-640-211A-328 (1-673)
Qy 17 LysGlyProTyrTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGly 36
Db 2 AGAGTCCGTGGAGCGGTGGAGGAGGACCTCACCTCGTCAATTACATTGCGCAACCGCGC 61
Qy 37 GluGlyValTyrAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCys 56
Db 62 GAAGGACGCTGGAACCTCCCTCGCGCGCAGCGAGGTTTGAACCGACCGGAAAGAGTTGC 121
Qy 57 ArgLeuArgTyrLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGlu 76
Db 122 CGCTCGCGTGGTGAATTAATACCTCCCGCGCGAGCTTCGGCGCGGAAACATAAACCTCGAA 181
Qy 77 GluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAla 96
Db 182 GAGCAGCTCTTGATCTCTCGAGCTCCATTCGCGTGGGGCAATCGATGGTGAAGATCGCG 241
Qy 97 LysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGln 116
Db 242 CAACACTTTCGGCGGAGGACCGCAACAGAGATCAAGAACTACTTGGCGAAACCGGGTGCAG 301
Qy 117 LysHisIleLysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsn 136
Db 302 AAGCAGCGCAAGCAGCTC-----AAATGTGACGTCAACAGCAGCAGTTCGAAG 349
Qy 137 AspHis-----GlnAlaSerThrSer 143
Db 350 GACGCCATGAATACCTCTGGATGCCGAGGCTGGTCGAGAGGATCCAAAGCGGCTCCGCC 409
Qy 144 HisValSerThr-----MetAlaGluProMetGluMetTyrSer 156
Db 410 TCTGTCTCGACCGCTACTGTGCGCGCGCCATGGCAGCCGCCACCAATGGCCACC 469
Qy 157 -----ProProCysTyrGlnGlyMetLeuGlu 165
Db 470 ACGCAGCATCCAACTACGCGCGCATGGCTTCCCGCGCGCCCTTCGGCGGCGATGGCGCGC 529
Qy 166 ProPheSer-----ThrGlnPheProThrIleAsnProAspGlnSerSerCys 181
Db 530 GACTTACGGGCGGGCGAGTGAATGTGGCCGCCAGCTACAGCACCCTCCGGAGAACTCTTGC 589
Qy 182 CysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMetGln 201
Db 590 ACGACG-----GCGTCTTCGACTCATTCGTTGCGCGCAG 622
Qy 202 Leu 202
Db 623 GTC 625

RESULT 5
US-09-640-211A-1934
; Sequence 1934, Application US/09640211A
; Patent No. 6833446
```

GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1934
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1934

Alignment Scores:
Pred. No.: 1,27e-52 Length: 524
Score: 478.00 Matches: 81
Percent Similarity: 84.87% Conservative: 20
Best Local Similarity: 68.07% Mismatches: 18
Query Match: 42.00% Indels: 0
DB: 4 Gaps: 0

US-10-659-869A-36 (1-206) x US-09-640-211A-1934 (1-524)

QY 3 LysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrpThrMet 22
DB 154 AGGAAGATGAGCATGTCGCGAGAAAGAGGGGTGACCTGCGAAGGGGCCCATGACTGCG 213
QY 23 GluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSer 42
DB 214 GAGGAAGCAATTTGCTCATTCATCGATCATCGCCAGCGGAGGAGCTGGAAATG 273
QY 43 LeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsn 62
DB 274 TTGGCGAAGCGCAGGATGGAAGAGAACTGGCAAAAGCTGCAGATTAAAGTGGCTGAAT 333
QY 63 TyrLeuArgProAspValArgGlyAsnIleThrProGluGlnLeuLeuIleMet 82
DB 334 TACCTGAGACCGGACATCAAGCGCGGAATCTCACCCGCAAGACACGCTCATGATCCTT 393
QY 83 GluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
DB 394 GAATTCACCAAAATGGGGCAACAGGTGTCGAAAATCGCGCAGTATCTCCAGGAAGG 453
QY 103 ThrAspAsnGluLeuLysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGln 121
DB 454 ACAGATAACGAGATCAAGAACTACTTGGAGGACGCGGTGCAGAAAGCGCGCCAG 510

RESULT 6

US-09-533-029-39
; Sequence 39, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Kiechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G239
US-09-533-029-39

Alignment Scores:

Pred. No.: 1,44e-51 Length: 1347
Score: 475.00 Matches: 92
Percent Similarity: 66.48% Conservative: 27
Best Local Similarity: 51.40% Mismatches: 36
Query Match: 41.74% Indels: 24
DB: 4 Gaps: 4

US-10-659-869A-36 (1-206) x US-09-533-029-39 (1-1347)

QY 11 GlnAspProGluValArgLysGlyProTrpThrMetGluGluAspLeuLeuMetAsn 30
DB 46 GAAGATTCTGATGTACGAAAGGTCCATGGACCGAGGAAGATGCAATCCTAGTCAAC 105
QY 31 TyrIleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLys 50
DB 106 TTCGTCTCTATTATGGGATGCTCGTTGGNACCACATCGCTCTCTCTGGGCTAAG 165
QY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArg 70
DB 166 CGAACTGTAAGAGCTTGTAGATTAAAGTGGCTTAATTACTTACGTCCAGATGTTAGAAGA 225
QY 71 GlyAsnIleThrProGluGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsn 90
DB 226 GGCAACATCACTCTCGAAGAACAAATTTATGATCTCTCAAACTCCATTCCTTTGGGCAAT 285
QY 91 ArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluLeuLysAsnTyr 110
DB 286 AGTGTTGCGAAGATTGGCAATATCTACCGGGAAGACAGATATGAATAAGAAATAT 345
QY 111 TrpArgThrArgIleGlnLysHisIleLysGlnAla-----GluAsnPhe 125
DB 346 TGGAGAACTCGAGTCCAAAGCAAGCAACACCTAAGATCGATGTTAAGTAATCTT 405
QY 126 GlnGlnSerSerAsnAsn-----SerGluLeuAsnAspHis 138
DB 406 TTCAAGGAGACTATGAGAAATGTTTGGATCCGAGATTAGTGAACGAAATCAACGCCAA 465
QY 139 GlnAlaSerThr-----SerHisValSerThrMetAlaGluProMet 152
DB 466 TCATTACCCACACGTGTGAACAAGTGGAGTCAATGATCACCAGCCCAAGTCAACCACTT 525
QY 153 GluMetTyrSerProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhe 171
DB 526 AACGAACCGAGTCCG-----GTCGAGCGCGGTTTCGTTCAATTC 564

RESULT 7

US-09-640-211A-283
; Sequence 283, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:

; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16

```

; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-283

Alignment Scores:
Pred. No.: 2,11e-45 Length: 389
Score: 421.00 Matches: 81
Percent Similarity: 88.17% Conservative: 1
Best Local Similarity: 87.10% Mismatches: 9
Query Match: 36.99% Indels: 2
DB: 4 Gaps: 1

US-10-659-869A-36 (1-206) x US-09-640-211A-283 (1-389)

Qy 1 MetAspLysLys-----GlnGlnCysLysThrSerGlnAspProGluValArgLysGly 18
Db 44 ATGGACAAGAAGCCAGACGACGAGTGGTAAAGTCCCAAGATGTGCGAGGTGAGAAAGGG 103

Qy 19 ProTrpThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
Db 104 CCGTGGACGATGGAAGAGGATCTCATCTCACTACATAGCAATACACGGGAAAGGC 163

Qy 39 ValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db 164 AGTTGGAACTCCCTAGCCAAAGCTGCTGCTTAAACGTAACCGGGAAGAGTTGTGGCTC 223

Qy 59 ArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGln 78
Db 224 CGGTGGCTGAATCTATCTGCGACCGGACGTCGCGAGAGGCAACATCACTACTGAGGAGCAG 283

Qy 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArg 91
Db 284 CTCCTGATCATGGAACATGTCATGCCAAGTGGGGAACAGG 322

RESULT 9
US-09-640-211A-483
; Sequence 483, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 483
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-483

Alignment Scores:
Pred. No.: 7,78e-44 Length: 516
Score: 410.50 Matches: 73
Percent Similarity: 82.08% Conservative: 14
Best Local Similarity: 68.87% Mismatches: 12
Query Match: 36.07% Indels: 7
DB: 4 Gaps: 1

US-10-659-869A-36 (1-206) x US-09-640-211A-483 (1-516)

Qy 2 AspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrpThr 21
Db 219 GATCAAGAGGAGCAATGT-----GTGAGGAGGGGACCTTGAGCT 257

Qy 22 MetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGlyValTrpAsn 41
Db 258 GTTGATGAGGACATGACCCCTTATTCGATGCGTAACACCCCGGGGTGAGGTGCGATGNAAC 317

Qy 42 SerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeu 61
Db 318 ACAGTAGCCAAAATTTCGCGGGCTAAAGAGAAACAGGAAGAGCTGCAGATTGAGATGGCTT 377

Qy 62 AsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeuIle 81
Db 378 AATTATCTTCGGCCCGATGTTAAACGTTGAAACATAACGCCGGAAGAGCAGCTATTATATC 437

Qy 82 MetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGly 101
Db 438 CTTGAACCTCCACCGTCTCTCGGGTAAACAGATTGTCACGCAACTCCAGGC 497

; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-283

Alignment Scores:
Pred. No.: 2,11e-45 Length: 389
Score: 421.00 Matches: 81
Percent Similarity: 88.17% Conservative: 1
Best Local Similarity: 87.10% Mismatches: 9
Query Match: 36.99% Indels: 2
DB: 4 Gaps: 1

US-10-659-869A-36 (1-206) x US-09-640-211A-283 (1-389)

Qy 1 MetAspLysLys-----GlnGlnCysLysThrSerGlnAspProGluValArgLysGly 18
Db 44 ATGGACAAGAAGCCAGACGACGAGTGGTAAAGTCCCAAGATGTGCGAGGTGAGAAAGGG 103

Qy 19 ProTrpThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
Db 104 CCGTGGACGATGGAAGAGGATCTCATCTCACTACATAGCAATACACGGGAAAGGC 163

Qy 39 ValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db 164 AGTTGGAACTCCCTAGCCAAAGCTGCTGCTTAAACGTAACCGGGAAGAGTTGTGGCTC 223

Qy 59 ArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGln 78
Db 224 CGGTGGCTGAATCTATCTGCGACCGGACGTCGCGAGAGGCAACATCACTACTGAGGAGCAG 283

Qy 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArg 91
Db 284 CTCCTGATCATGGAACATGTCATGCCAAGTGGGGAACAGG 322

RESULT 8
US-09-640-211A-1496
; Sequence 1496, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1496
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1496

Alignment Scores:
Pred. No.: 2,36e-45 Length: 417
Score: 421.00 Matches: 81
Percent Similarity: 88.17% Conservative: 1
Best Local Similarity: 87.10% Mismatches: 9
Query Match: 36.99% Indels: 2
DB: 4 Gaps: 1

US-10-659-869A-36 (1-206) x US-09-640-211A-1496 (1-417)

```

```
Qy 102 ArgThrAspAsnGluile 107
Db 498 AGGACTGACAAACGAAATC 515

RESULT 10
US-09-640-211A-588
; Sequence 588, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 588
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-588

Alignment Scores:
Pred. No.: 7,78e-44 Length: 516
Score: 410.50 Matches: 73
Percent Similarity: 82.08% Conservative: 14
Best Local Similarity: 68.87% Mismatches: 12
Query Match: 36.07% Indels: 7
Gaps: 4

US-10-659-869A-36 (1-206) x US-09-640-211A-588 (1-516)
Qy 2 AspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProThr 21
Db 219 GATCAAGAGGAGCAATGT- - - - - GTGAGGAGGGGACCTTGGACT 257

Qy 22 MetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsn 41
Db 258 GTTGATGAGGACATGAGCGCTTATTGCGTACCGCCCGGGTGAAGGTCGATGGAAC 317

Qy 42 SerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeu 61
Db 318 ACAGTAGCCAAATTTGCGAGGCTAAAGAGACAGGAAAGAGCTGCAGATTGAGATGGCTT 377

Qy 62 AsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeuile 81
Db 378 AATTATCTCGGCCCGGATGTTAAACGTGGAAACATAACCGCCGGAAGAGCAGCTATTAAATC 437

Qy 82 MetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGly 101
Db 438 CTTGAACCTCCACGCTCTCTGGGGTAACAGATGGTCCAAAGATTGCGCAACACTCCCGGC 497

Qy 102 ArgThrAspAsnGluile 107
Db 498 AGGACTGACAAACGAAATC 515

RESULT 11
US-09-640-211A-565
; Sequence 565, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A

US-10-659-869A-36 (1-206) x US-09-640-211A-565 (1-542)
Qy 12 AspProGluValArgLysGlyProThrMetGluGluAspLeuIleLeuMetAsnTyr 31
Db 81 GAAGCCGATACGAGGAAGGGTCCATGCATCTGTGAAGAGGACATGCAGCTT---GGTATT 137

Qy 32 IleAlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArg 51
Db 138 GTAAATTTGACGAGGAAGGACGCTGGAACCTTTCTGCCAGAGCATCTGGCCTCCAGAGA 197

Qy 52 AsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgGly 71
Db 198 ACTGGTAAGAGCTCCCGCTTAGGTGGTTAACTATCTCCGCGCTGATCTCAAGCGGAGC 257

Qy 72 AsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArg 91
Db 258 AAGATCACTCTGAAGAAGAACGTTTGATTATTGAACTCCCATCGCGTGGGGAATAAGG 317

Qy 92 TrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrp 111
Db 318 TGGTCTCGTATTGCACAAAGTTTACCGGGAAGGACGGAACAATGAATCAAGATTTCTGG 377

Qy 112 ArgThrArgIleGln 116
Db 378 AGAACTCGTATGAAG 392

RESULT 12
US-09-640-211A-1559
; Sequence 1559, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1559
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1559

Alignment Scores:
Pred. No.: 2,98e-43 Length: 389
Score: 404.50 Matches: 70
Percent Similarity: 84.62% Conservative: 18
Best Local Similarity: 67.31% Mismatches: 15
Query Match: 35.54% Indels: 1
Gaps: 4
```



```

299  CCTGACGAGGAAGAGCTTATCATCAAACTCCACAAACTCTTGGGAAAACAGATGGTCTTTG 355
QY      95  IleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArg 114
Db      359  ATAGCTGGGAGGCTTCCAGGACCAACAGACAAATGAAATAAAGAATTACTGGAACACCAAC 418
QY      115  IleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnGlnSerSerAsnAsnSerGlu 134
Db      419  TTAAGTAAAGAAGAGTTCTCCGATCGTCAAAAGACTCACCCCGCTCTTCGAAAAATCCCGAG 478
QY      135  Ile-----AsnAspHisGlnAlaSer 141
Db      479  GCGGCTCGACGAGAACTGCTGTAATGCGCAATACCAATGGTAATGGTAGTGGTAGTTCC 538
QY      142  ThrSerHisVal-----SerThrMetAlaGluProMetGluMetTyrSerPro 157
Db      539  TCGACACACGTGTGCGGACGAAGGGCGCAAGGGTCTCCAAGGTT---TTCATAAACCT 595
QY      158  ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAsp 177
Db      596  CCTCACTAC-----ACAAA-----AACAGAGACCCCAAG 625

RESULT 14
US-09-533-029-85
; Sequence 85, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Onaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G229
US-09-533-029-85

Alignment Scores:
Pred. No.: 1,53e-41 Length: 1209
Score: 397.50 Matches: 71
Percent Similarity: 77.05% Conservative: 23
Best Local Similarity: 58.20% Mismatches: 25
Query Match: 34.93% Indels: 3
DB: 4 Gaps: 1

US-10-659-869A-36 (1-206) x US-09-533-029-85 (1-1209)

QY      15  ValArgLysGlyProThrThrMetGluLysAspLeuIleLeuMetAsnTyrIleAlaAsn 34
Db      74  ATCAAGAGGGGGCGGTGGACGGCGGGAGGAGCAGATTCTCTCAACTACATTCAATCC 133

```

```
Qy 35 HisGlyGluGlyValTrpAsnSerLeuAlaLysAlaLysAlaLysAlaLysAlaLysAlaLys 54
   ::::::::::::::::::::
Db 134 AATGGTGAAGGTTCTTGGAGATCTCTCCCAAAATGCGGATTAAAGAGGTGGAAAG 193
   ::::::::::::::::::::

Qy 55 SerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValaArgArgGlyAsnIleThr 74
   ::::::::::::::::::::
Db 194 AGCTGTAGATTGAGATGATAAATCTATCTAAGATCAGACCTCAAGCGTGGAAACATAACT 253
   ::::::::::::::::::::

Qy 75 ProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLys 94
   ::::::::::::::::::::
Db 254 CCAGAAAGAAAGAACTCGTTGTTAAATTCATTCCACTTTGGGAAACAGGTGCTCACTA 313
   ::::::::::::::::::::

Qy 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArg 114
   ::::::::::::::::::::
Db 314 ATCGGGGTCTATCTACCGAGGAGAGACAGACACGAAATATAAATATTGGAACTCTCAT 373
   ::::::::::::::::::::

Qy 115 IleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGlu 134
   ::::::::::::::::::::
Db 374 CTCAGCCGTAAACTC-----CACAACTTCATTAGGAAGCCATCCATCTCTCAAGAC 424
   ::::::::::::::::::::

Qy 135 IleAsn 136
   ::::::::::::::
Db 425 GTCTCC 430
```

RESULT 15

US-09-453-387A-1
; Sequence 1, Application US/09453387A

; Patent No. 6828476

; GENERAL INFORMATION:

; APPLICANT: Wilkins, Thea A.

; TITLE OF INVENTION: The Regents of the University of California

; FILE OF INVENTION: Cotton Transcription Factors and Their Uses

; FILE REFERENCE: 023070-095600US

; CURRENT APPLICATION NUMBER: US/09/453,387A

; CURRENT FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 1006

; TYPE: DNA

; ORGANISM: Gossypium hirsutum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (59)...(943)

; OTHER INFORMATION: GIMYBI

US-09-453-387A-1

Alignment Scores:

Pred. No.:	1.53e-41	Length:	1006
Score:	396.50	Matches:	89
Percent Similarity:	52.89%	Conservative:	30
Best Local Similarity:	39.56%	Mismatches:	80
Query Match:	34.84%	Indels:	26
DB:	4	Gaps:	4

US-10-659-869A-36 (1-206) x US-09-453-387A-1 (1-1006)

```
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
   ||| ::|
Db 59 ATGGGACGATCACCTTGTGTGAAAGGCT-----CATACCACAAAGGTGCCTGG 109
   ||| |||

Qy 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
   ||| |||
Db 110 ACCAAGAGAGAGATCAACGCCCTCATCACTACATCCGTGTCCTGTTGAAGGCTGCTGG 169
   ||| |||

Qy 41 AsnSerLeuAlaLysAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
   ||| |||
Db 170 CGTTCCTCCCAAGAGTCTGGGCTGCTTAGATGCTGGTAAAGATTGAGATTAAAGATGG 229
   ||| |||

Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeu 80
   ||| |||
Db 230 ATAAACTACTTGAGGCCCTGATCTTAAAGAGAGAGAAATTTCACTGAAGAAAGATGAGCTT 289
   ||| |||
```

```
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
   ||| ::|
Db 290 ATCATCAAGCTTCACATTTACTTGGAAACAAATGGTTCATGTTGCTGGAGATTACCA 349
   ||| |||

Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIle--- 119
   ||| |||
Db 350 GGAAGAACAGATATGAGATAAAGAACTACTTGGAAACACACACATCAAAAGAAAGCTTATA 409
   ||| |||

Qy 120 -----LysGlnAlaGluAsnPheGlnGlnSerSerAsnAsnSer--- 133
   ||| |||
Db 410 AGCAGAGGAATTGATCCACAAACTCATCGTCCTCTCAATCAACGCGCAATACCAACACA 469
   ||| |||

Qy 134 -----GluIleAsnAspHisGlnAlaSerThrSerHisValSer 146
   ||| |||
Db 470 GTCACAGCCCCCAGGAAATGGATTTTCAAGAACTCGCCACATCCGTTTCCAAATCCAGT 529
   ||| |||

Qy 147 ThrMetAlaGluProMetGluMetTyrSerProCysTyrGlnGlyMetLeuGluPro 166
   ||| |||
Db 530 TCCATCAAAAACCCGTCCTCTGGATTTCATTAATGAATTTCAATTCAGTCCCAACACA 589
   ||| |||

Qy 167 PheSerThrGln-PheProThrIleAsnPro-----AspG1 178
   ||| |||
Db 590 GATTCCCTTGAAGAACCCCAACTGTACAGCCAGCAGTGGCATGACTACAGATGAAGAGCAA 649
   ||| |||

Qy 178 nSerSerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTr 198
   ||| |||
Db 650 CAAGAACAGCTGCACAGAACAGCAGCAATACCGTCCGAGCAATGGGCAAGACATAAATTG 709
   ||| |||

Qy 198 pSerMetGlnLeu 202
   ||| |||
Db 710 GAGCTGTCGATTG 722
   ||| |||
```

Search completed: September 28, 2005, 16:08:00

Job time : 179 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2005, 15:03:12 ; Search time 639 Seconds
(without alignments)
2155.513 Million cell updates/sec

Title: US-10-659-869A-36

Perfect score: 1138

Sequence: 1 MDKQKQCKTQDPEVRKGPW.....NNINYSMEDSWSQLNGD 206

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7400732 seqs, 3343137571 residues

Total number of hits satisfying chosen parameters: 14801464

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO_spool/US10659869/runat_26092005_143448_11431/app_query.fasta_1.391

-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62

-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPSTR=500 -MINLEN=0

-MAXLEN=2000000000 -WAPR=US10659869@cgn_1_1.480 @runat_26092005_143448_11431

-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQRY -NEG SCORES=0 -WAIT -DSBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubna/US10G_PUBCOMB.seq*
- 20: /cgn2_6/ptodata/1/pubna/US10H_PUBCOMB.seq*
- 21: /cgn2_6/ptodata/1/pubna/US10I_PUBCOMB.seq*
- 22: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq*
- 23: /cgn2_6/ptodata/1/pubna/US11A_PUBCOMB.seq*
- 24: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq*
- 25: /cgn2_6/ptodata/1/pubna/US16_NEW_PUB.seq*
- 26: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1138	100.0	863	14	US-10-021-811-35
2	1138	100.0	863	18	US-10-659-869-35
3	1138	100.0	947	17	US-10-425-114-10661
4	1092	96.0	968	17	US-10-374-780A-1537
5	1092	96.0	968	18	US-10-412-699B-1631
6	1062	93.3	988	14	US-10-021-811-29
7	1062	93.3	988	18	US-10-659-869-29
8	930	81.7	1097	18	US-10-424-599-85949
9	889.5	78.2	988	17	US-10-374-780A-1536
10	889.5	78.2	988	18	US-10-412-699B-1630
11	885	77.8	877	18	US-10-425-114-10695
12	885	77.8	1094	18	US-10-425-114-10688
13	874.5	76.8	805	14	US-10-021-811-37
14	874.5	76.8	805	18	US-10-659-869-37
15	829.5	72.9	910	14	US-10-021-811-33
16	829.5	72.9	910	18	US-10-659-869-33
17	810	71.2	969	17	US-10-374-780A-1538
18	810	71.2	969	18	US-10-412-699B-1632
19	760	66.8	681	21	US-10-487-901-3444
20	759.5	66.7	862	20	US-10-856-499-2076
21	724	63.6	681	17	US-10-341-961A-197
22	663	58.3	731	17	US-10-374-780A-1981
23	663	58.3	731	18	US-10-412-699B-173
24	663	58.3	731	18	US-10-412-699B-1757
25	653.5	57.4	1033	15	US-09-533-029-87
26	653.5	57.4	1033	15	US-10-278-536-85
27	653.5	57.4	1033	17	US-10-225-066A-451
28	653.5	57.4	1033	17	US-10-374-780A-1969
29	653.5	57.4	1033	18	US-10-412-699B-1745
30	653.5	57.4	1033	22	US-10-225-066A-451
31	638.5	56.1	462	21	US-10-487-901-6907
32	625.5	55.0	907	20	US-10-856-499-2077
33	592	52.0	1011	17	US-10-225-068-1
34	592	52.0	1011	17	US-10-225-068-21
35	592	52.0	1011	17	US-10-225-066A-229
36	592	52.0	1011	17	US-10-302-267-129
37	592	52.0	1011	17	US-10-225-067-97
38	592	52.0	1011	17	US-10-374-780A-257
39	592	52.0	1011	18	US-10-412-699B-841
40	592	52.0	1011	21	US-10-225-068-1
41	592	52.0	1011	21	US-10-225-068-21
42	592	52.0	1011	22	US-10-225-066A-229
43	591	51.9	530	14	US-10-021-811-31
44	591	51.9	530	18	US-10-659-869-31
45	591	51.9	530	18	Sequence 35, Appl
					Sequence 35, Appl
					Sequence 10661, A
					Sequence 1537, Ap
					Sequence 1631, Ap
					Sequence 29, Appl
					Sequence 29, Appl
					Sequence 85949, A
					Sequence 1536, Ap
					Sequence 1630, Ap
					Sequence 10695, A
					Sequence 10688, A
					Sequence 126212,
					Sequence 37, Appl
					Sequence 37, Appl
					Sequence 33, Appl
					Sequence 1538, Ap
					Sequence 1632, Ap
					Sequence 3444, Ap
					Sequence 2076, Ap
					Sequence 197, App
					Sequence 1757, Ap
					Sequence 87, Appl
					Sequence 451, App
					Sequence 1969, Ap
					Sequence 1745, App
					Sequence 451, App
					Sequence 6907, Ap
					Sequence 2077, Ap
					Sequence 1, Appl
					Sequence 21, Appl
					Sequence 229, App
					Sequence 129, App
					Sequence 97, Appl
					Sequence 257, App
					Sequence 841, App
					Sequence 1, Appl
					Sequence 21, Appl
					Sequence 229, App
					Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-10-021-811-35

; Sequence 35, Application US/10021811

; Publication No. US2003004007A1

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.

; APPLICANT: Fang, Yiwen

; APPLICANT: Odell, Joan

; APPLICANT: Weng, Zude

; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs

; FILE REFERENCE: BB1294 US NA

; CURRENT APPLICATION NUMBER: US/10/021.811

; CURRENT FILING DATE: 2001-12-14

; PRIOR APPLICATION NUMBER: 60/110,609

; PRIOR FILING DATE: 1998-December-02

; NUMBER OF SEQ ID NOS: 63

```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Glycine max
US-10-021-811-35

Alignment Scores:
  Pred. No.:      1,62e-135      Length:      863
  Score:          1138.00        Matches:     206
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:      100.00%      Indels:       0
  DB:               14          Gaps:         0

US-10-659-869A-36 (1-206) x US-10-021-811-35 (1-863)

Qy  1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db  29 ATGGATAAAAAACACACAGTGTAAAGACGCTCTCAAGATCTCTGAAGTGAGAAAAAGGCGCTTGG 88
Qy  21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db  89 ACAATGGGAAGAAGACTTGTATCTTGATGAACATATATTGCAAAATCATGGGGAAGGTGTTTGG 148
Qy  41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db  149 AACTCTTTGGCCAAAGCTGTGCTCTCAACGTAACGGAAGAAGATTGCCGCTTAAGGTGG 208
Qy  61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db  209 CTAAATTACCTCCGTCCTGATGTTAGAGAAGGGAATATTACACCCGAGGAACAACCTTTTG 268
Qy  81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db  269 ATTATGGAGCTCCACGCAAGTGGGGAACACAGGTGGTCCAAAATTGCCAAGCATCTACCT 328
Qy  101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db  329 GGAAGGACTGATATGAGATCAAGAACTATTTGGAGGACAAGATCCAGAACACATCAAG 388
Qy  121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db  389 CAAGCTGAGAACTTTTCAGCAACACAGATAGTAATAATCTGAGATAAATGATCACCAGCT 448
Qy  141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
Db  449 AGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTATTCTCCACCTGTTAT 508
Qy  161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db  509 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTCCATCAANTTAATCTGTATCAATCCAGT 568
Qy  181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db  569 TGTGTACCAATGACAAACAACATTAACATTATTGGAGCATGGAGATAGCTGGTCAATG 628
Qy  201 GlnLeuLeuAsnGlyAsp 206
Db  629 CAATTACTGAACGGTGAT 646

RESULT 2
US-10-659-869-35
; Sequence 35, Application US/10659869
; Publication No. US20040040057A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/659,869

; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Glycine max
US-10-659-869-35

Alignment Scores:
  Pred. No.:      1,62e-135      Length:      863
  Score:          1138.00        Matches:     206
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:      100.00%      Indels:       0
  DB:               18          Gaps:         0

US-10-659-869A-36 (1-206) x US-10-659-869-35 (1-863)

Qy  1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db  29 ATGGATAAAAAACACACAGTGTAAAGACGCTCTCAAGATCTCTGAAGTGAGAAAAAGGCGCTTGG 88
Qy  21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db  89 ACAATGGGAAGAAGACTTGTATCTTGATGAACATATATTGCAAAATCATGGGGAAGGTGTTTGG 148
Qy  41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db  149 AACTCTTTGGCCAAAGCTGTGCTCTCAACGTAACGGAAGAAGATTGCCGCTTAAGGTGG 208
Qy  61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db  209 CTAAATTACCTCCGTCCTGATGTTAGAGAAGGGAATATTACACCCGAGGAACAACCTTTTG 268
Qy  81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db  269 ATTATGGAGCTCCACGCAAGTGGGGAACACAGGTGGTCCAAAATTGCCAAGCATCTACCT 328
Qy  101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db  329 GGAAGGACTGATATGAGATCAAGAACTATTTGGAGGACAAGATCCAGAACACATCAAG 388
Qy  121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db  389 CAAGCTGAGAACTTTTCAGCAACACAGATAGTAATAATCTGAGATAAATGATCACCAGCT 448
Qy  141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
Db  449 AGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGTATTCTCCACCTGTTAT 508
Qy  161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db  509 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCTCCATCAANTTAATCTGTATCAATCCAGT 568
Qy  181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db  569 TGTGTACCAATGACAAACAACATTAACATTATTGGAGCATGGAGATAGCTGGTCAATG 628
Qy  201 GlnLeuLeuAsnGlyAsp 206
Db  629 CAATTACTGAACGGTGAT 646

RESULT 3
US-10-425-114-10661
; Sequence 10661, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
```

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 10661
LENGTH: 947
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700957658_FLI
US-10-425-114-10661

Alignment Scores:
Pred. No.: 1.85e-135 Length: 947
Score: 1138.00 Matches: 206
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-659-869A-36 (1-206) x US-10-425-114-10661 (1-947)

Qy 1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
Db 48 ATGGATAAAAACACAGTGTGAAGACGTCTCAAGATCTCTGAAGTGAGAAAAGGCCCTTGG 107
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40
Db 108 ACAATGGAGAGACTTGAATCTTGATGACTATATTGCAATCATATGGGAGGTGTTGG 167
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
Db 168 AACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGGAAGAGTTGCGGCTAAGGTGG 227
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 228 CTAATTAATCTCCCTCGTCTGATGTTAGAGAGGGGAATATTACCCCGAGGAACTTTTG 287
Qy 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
Db 288 ATTATGAGGTCTCCAGCAAGTGGGGAAACAGGTGGTCCAAATTTGCCAAGCATCTACCT 347
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrrpArgThrArgIleGlnLysHisIleLys 120
Db 348 GGAAGGACTGATAATGAGATCAAGAACTATTGGAGGACAGGATCCAGAGCATCAAG 407
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 408 CAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTTCTGAGATAAATGATCAACAGCT 467
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 468 AGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGATATTCTCCACCTGTTAT 527
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 528 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTACAAATTAATCTGTGATCAATCAGT 587
Qy 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTrrpSerMetGluAspSerTrrpSerMet 200
Db 588 TGTGTGACCAATGACAAACAACTTAATTTGGAGCATGGAGGATAGTGTGTCATG 647
Qy 201 GlnLeuLeuAsnGlyAsp 206
Db 648 CAATTACTGAAACGGTGAT 665

RESULT 4
US-10-374-780A-1537
Sequence 1537, Application US/10374780A
Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Mareha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2306
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1537
LENGTH: 968
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1322
US-10-374-780A-1537

Alignment Scores:
Pred. No.: 1.53e-129 Length: 968
Score: 1092.00 Matches: 198
Percent Similarity: 97.09% Conservative: 2
Best Local Similarity: 96.12% Mismatches: 6
Query Match: 95.96% Indels: 0
DB: 17 Gaps: 0

US-10-659-869A-36 (1-206) x US-10-374-780A-1537 (1-968)

Qy 1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
Db 72 ATGGATAAAAACACAGTGTGCAACACGTCTCAAGATCTCTGAAGTGAGAAAAGGCCCTTGG 131
Qy 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGlyValTrrp 40
Db 132 ACAATGGAGAGACTTGAATCTTGATCAACTATATTGCAAAATCATGGGGAAGGTGTTGG 191
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
Db 192 AACTCTTTGGCCAAAGCTGCTGCTCAACAGTACCGAAGAGTGTGCCGCTAAGGTGG 251

```
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 252 CTAAACTACCTCCGCTCTGATGTTAGAGAGGGAAATATTACCCCGAGGAACAACATTTTG 311
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 312 ATCATGGAGCTTCAGCAAGAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGCATCTACCT 371
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 372 GGAAGGACTGATAATGAGATCAAGAACTATTGGAGGCAAGGATCCAGAACATCAAG 431
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 432 CAAGCTGAGAACTTTCACAAACAGAGTAGTAATTAATCTGAGATAAATGATCCCAAGCT 491
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 492 TGCATAGCCATGTGTCCACCATGGCGCAGCCCATGGAGACCTATTCTCCACCCCTGTTAT 551
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 552 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTACAATTAATCTCCTGATCAATCCAGT 611
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db 612 TGTGTACCAATGACAAACAAACAACTAATTATTTGGAGCATGGAGATAGCTGGTCAATG 671
Qy 201 GlnLeuLeuAsnGlyAsp 206
Db 672 CAATTACTGAACGGTGAT 689
```

RESULT 5

```
US-10-412-699B-1631
; Sequence 1631, Application US/10412699B
; Publication No. US20040045049A1
```

GENERAL INFORMATION:

```
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1631
; LENGTH: 968
; TYPE: DNA
; ORGANISM: Glycine max
US-10-412-699B-1631
```

Alignment Scores:

```
Pred. No.: 1.53e-129 Length: 968
Score: 1092.00 Matches: 198
Percent Similarity: 97.09% Conservative: 2
Best Local Similarity: 96.12% Mismatches: 6
Query Match: 95.96% Indels: 0
DB: 18 Gaps: 0
```

US-10-659-869A-36 (1-206) x US-10-412-699B-1631 (1-968)

```
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 72 ATGGATAAAAAACAACTGTGCAACACGTCTCAAGATCTCTGAAGTGAGAAAAAGGCGCTTGG 131
Qy 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db 132 ACAATGGAGAGACTTGATCTTGATCACTATATTGCAATATCATGGGAAGGTGTTGG 191
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 192 AACTCTTTGGCCAAAGCTGCTGCTCAACAGCTACCGGAAAGAGTTGCGCGCTAAGGTGG 251
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 252 CTAAACTACCTCCGCTCTGATGTTAGAGAGGGAAATATTACCCCGAGGAACAACATTTTG 311
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 312 ATCATGGAGCTTCAGCAAGAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGCATCTACCT 371
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 372 GGAAGGACTGATAATGAGATCAAGAACTATTGGAGGCAAGGATCCAGAACATCAAG 431
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 432 CAAGCTGAGAACTTTCAGCAACAGAGTAGTAATTAATCTGAGATAAATGATCCCAAGCT 491
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 492 TGCATAGCCATGTGTCCACCATGGCGCAGCCCATGGAGACCTATTCTCCACCCCTGTTAT 551
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 552 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTCCCTACAATTAATCTCCTGATCAATCCAGT 611
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db 612 TGTGTACCAATGACAAACAAACAACTAATTATTTGGAGCATGGAGATAGCTGGTCAATG 671
Qy 201 GlnLeuLeuAsnGlyAsp 206
Db 672 CAATTACTGAACGGTGAT 689
```

RESULT 6

US-10-021-811-29

```
; Sequence 29, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
US-10-021-811-29

Alignment Scores:
Pred. No.: 1,11e-125 Length: 988
Score: 1062.00 Matches: 197
Percent Similarity: 95.67% Conservative: 2
Best Local Similarity: 94.71% Mismatches: 7
Query Match: 93.32% Indels: 2
DB: 14 Gaps: 2

US-10-659-869A-36 (1-206) x US-10-021-811-29 (1-988)
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTirp 20
Db 61 ATGGATAAAAAACAACGTGTGCAACACGTCTCAAGATCTCTGAAGTGAGAAAAAGGACCTTGG 120
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTirp 40
Db 121 ACGATGAAGAAGACTTGTATCTTGATCAACTATATTTGCAAAATCATATGGGGAAGGTGTTGG 180
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlnGlyLysSerCysArgLeuArgTirp 60
Db 181 AATTCTTTGGCCAAAGCTGTGCTCAACGTACCGGAAGAGTTCGCCGCTAAGGTGG 240
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 241 CTAAACTACCTCCGTCCTGATGTTAGAGAGGGAATATTACACCCGAGGAACTTTTG 300
Qy 81 IleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisLeuPro 100
Db 301 ATCATGGAGCTTACGCAAAAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGCATCTACCT 360
Qy 101 GlyArgThrAspAsnGluLysAsnTyrTirpArgThrArgIleGlnLysHisIleLys 120
Db 361 GGTAGGACAGATATAGATCAAGACTATTGGAGGACAGGATCCAGAGCATCATCAAG 420
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluLysLeuAsnAspHisGlnAla 140
Db 421 CAAGCTGAGAACTTTCAGCAACAATCAGCAATACTCTGAGATAAATGATCACAAGCT 480
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 481 AGCAGTAGCATGTTTCTACCATGGCTGAACCCATGGAGACCTATTCTCCACCCCTTTAT 540
Qy 161 GlnGlyMetLeuGluProPheSerThr---GlnPheProThrIleAsnProAspGlnSer 179
Db 541 CAAGGAATGTTAGAGCAATTTCTTCAATTCAGTTCCTCCCAATTAATCTTGATCAATCC 600
Qy 180 SerCysCysThrAsnAspAsnAsn---IleAsnTyrTirpSerMetGluAspSerTirp 198
Db 601 AGTTGTTGACCATGACAAACAACAGCATTAACTATTGGAGCATGGAGGATATCTGG 660
Qy 199 SerMetGlnLeuLeuAsnGlyAsp 206
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 661 TCAATGCAGTTACTGAACGGGGAT 684
RESULT 7
US-10-659-869-29
; Sequence 29, Application US/10659869
; Publication No. US20040040057A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/659,869
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
US-10-659-869-29

Alignment Scores:
Pred. No.: 1,11e-125 Length: 988
Score: 1062.00 Matches: 197
Percent Similarity: 95.67% Conservative: 2
Best Local Similarity: 94.71% Mismatches: 7
Query Match: 93.32% Indels: 2
DB: 18 Gaps: 2

US-10-659-869A-36 (1-206) x US-10-659-869-29 (1-988)
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTirp 20
Db 61 ATGGATAAAAAACAACGTGTGCAACACGTCTCAAGATCTCTGAAGTGAGAAAAAGGACCTTGG 120
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTirp 40
Db 121 ACGATGAAGAAGACTTGTATCTTGATCAACTATATTTGCAAAATCATATGGGGAAGGTGTTGG 180
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlnGlyLysSerCysArgLeuArgTirp 60
Db 181 AATTCTTTGGCCAAAGCTGTGCTCAACGTACCGGAAGAGTTCGCCGCTAAGGTGG 240
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 241 CTAAACTACCTCCGTCCTGATGTTAGAGAGGGAATATTACACCCGAGGAACTTTTG 300
Qy 81 IleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisLeuPro 100
Db 301 ATCATGGAGCTTACGCAAAAGTGGGAAACAGGTGGTCCAAAATTTGCCAAGCATCTACCT 360
Qy 101 GlyArgThrAspAsnGluLysAsnTyrTirpArgThrArgIleGlnLysHisIleLys 120
Db 361 GGTAGGACAGATATAGATCAAGACTATTGGAGGACAGGATCCAGAGCATCATCAAG 420
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluLysLeuAsnAspHisGlnAla 140
Db 421 CAAGCTGAGAACTTTCAGCAACAATCAGCAATACTCTGAGATAAATGATCACAAGCT 480
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 481 AGCAGTAGCATGTTTCTACCATGGCTGAACCCATGGAGACCTATTCTCCACCCCTTTAT 540
Qy 161 GlnGlyMetLeuGluProPheSerThr---GlnPheProThrIleAsnProAspGlnSer 179
Db 541 CAAGGAATGTTAGAGCAATTTCTTCAATTCAGTTCCTCCCAATTAATCTTGATCAATCC 600
```

QY 180 SerCysCysThrAsnAspAsnAsnAen---lleAsnTyrTrpSerMetGluAspSerTrp 198
Db 601 AGTTGTTGTACCAATGACAAACAACAGCATTAATTTGGAGATCGGATATCTGG 660

QY 199 SerMetGlnLeuLeuAsnGlyAsp 206
Db 661 TCAATGCAGTTACTGAACGGGGAT 684

RESULT 8

US-10-424-599-85949
; Sequence 85949, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85949
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48625C.1
US-10-424-599-85949

Alignment Scores:

Pred. No.: 1.13e-108 Length: 1097
Score: 930.00 Matches: 175
Percent Similarity: 87.75% Conservative: 4
Best Local Similarity: 85.78% Mismatches: 9
Query Match: 81.72% Indels: 17
DB: 18 Gaps: 2

US-10-659-869A-36 (1-206) x US-10-424-599-85949 (1-1097)

QY 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 70 ATGGATAAAAAACAACAGTGTAAAGACGTCCTCAAGATCCTGAAGTGAAGAAAGGCGCTTGG 129

QY 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluValTrp 40
Db 130 ACAATGGGAAGAGACTTGATCTTGATGAACATATATGCAAAATCATGGGAAGGTGTTTGG 189

QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 190 AACTCTTTGGCCAAAGCTGCTGGTCTCAACGTAACGGAAAGAGTTGCCGCTAAGGTGG 249

QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeu 80
Db 250 CTAATACTACCTCCGCTCGATGTTAGAAGAGGGAATATTACACCCGAGGAACAACATTTTG 309

QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 310 ATCATGGAGCTCCAGCAAGTGGGGAACAGGTGCTCAAAATTTGCCAAGCATCTACCT 369

QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 370 GGTAGCAGACATAATGAGATCAAGAACTATTTCGAGGACAAGGATCCAGAGCACATCAAG 429

QY 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 430 CAAGCTGAGAACTTTTCAGCAACAAAA-----455

QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
Db 456 -----GTGTCCACCATGGCGCAACCCATCGGAACCTATTCTTCCACCCAGTTAT 503

QY 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 504 CAAGGAATGTTAGATCCATTTTCAATTCAGTTCCCAACA---AATCCTCATCATCTTAGT 560

QY 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrpSerMet 200
Db 561 TGTGTATCCATGACGACGACACAACAACACTATTTCGAGCATGGAGGATATCTGGTCAATG 620

QY 201 GlnLeuLeuAen 204
Db 621 CAATTAGCCAAT 632

RESULT 9

US-10-374-780A-1536
; Sequence 1536, Application US/10374780A
; Publication No. US2004001927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1536
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1322
US-10-374-780A-1536

Alignment Scores:
Pred. No.: 1.53e-103 Length: 988
Score: 889.50 Matches: 171
Percent Similarity: 85.65% Conservative: 8
Best Local Similarity: 81.82% Mismatches: 23
Query Match: 78.16% Indels: 7
DB: 17 Gaps: 5

US-10-659-869A-36 (1-206) x US-10-374-780A-1536 (1-988)


```
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 63 ATGGACAAAAA---CCATGCAACTCATCTCATGATCTCTGAAGTCAGAAAGGGACCATGG 119
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db 120 ACCATGGGAAGAACTTGATCTTGATAAATATATTTGCAATACCGGTGAAGGTGTTGG 179
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 180 AACTCTTTAGCCAAAGCTTCTGGTCTTAAACGAACGGGAAAGAGTTGTGCACTCGGTTGG 239
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 240 CTAACACTACCTTCGTCTGATGTTAGNAGAGGAACACATTACCCCGAGGAACAGCTTTTG 299
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 300 ATCATAGAACTTCATGCAAGTGGGCAATAGGTGGTCCAAAATTCGAAAGCATCTTTCCA 359
Qy 101 GlyArgThrAspAsnGluLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 360 GGAAGACTGACATGAGATTAAGAACTTCTGGAGAACAGGATCCAAAAGCACATTAAAG 419
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 420 CAAGCTGAGACTTCACAAACACAT-----GGTAATTCAGAGAATAATGATCATCAAGCA 473
Qy 141 -----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerPro 158
Db 474 AGCACTAGTACTAGCAAGTTCACCAATGGCATCCAAATGAGACTTCTCTCCACCC 533
Qy 159 CysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGln 178
Db 534 TCATACCAAGAACTTTGAGCCATTCCAACCTCAATTCCTTACATCACT--GATCAA 590
Qy 179 SerSerCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp 198
Db 591 TCAAGTTGTTGTACCAACCACCAACGACAAACAACTATTGAGCATCGAGGATATCTGG 650
Qy 199 ---SerMetGlnLeuLeuAsnGlyAsp 206
Db 651 TCGTCTATGCAATTAATCAATGGAGAT 677
```

RESULT 10

```
US-10-412-699B-1630
; Sequence 1630, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
```

```
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1630
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
US-10-412-699B-1630
```

Alignment Scores:

Pred. No.:	1.53e-103	Length:	988
Score:	889.50	Matches:	171
Percent Similarity:	85.65%	Conservative:	8
Best Local Similarity:	81.82%	Mismatches:	23
Query Match:	78.16%	Indels:	7
DB:	18	Gaps:	5

US-10-659-869A-36 (1-206) x US-10-412-699B-1630 (1-988)

```
Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 63 ATGGACAAAAA---CCATGCAACTCATCTCATGATCTCTGAAGTCAGAAAGGGACCATGG 119
Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db 120 ACCATGGGAAGAACTTGATCTTGATAAATATATTTGCAATACCGGTGAAGGTGTTGG 179
Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 180 AACTCTTTAGCCAAAGCTTCTGGTCTTAAACGAACGGGAAAGAGTTGTGCACTCGGTTGG 239
Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 240 CTAACACTACCTTCGTCTGATGTTAGNAGAGGAACACATTACCCCGAGGAACAGCTTTTG 299
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 300 ATCATAGAACTTCATGCAAGTGGGCAATAGGTGGTCCAAAATTCGAAAGCATCTTTCCA 359
Qy 101 GlyArgThrAspAsnGluLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 360 GGAAGAACTGACATGAGATTAAGAACTTCTGGAGAACAGGATCCAAAAGCACATTAAAG 419
Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 420 CAAGCTGAGACTTCACAAACACAT-----GGTAATTCAGAGAATAATGATCATCAAGCA 473
Qy 141 -----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerPro 158
Db 474 AGCACTAGTACTAGCAAGTTCACCAATGGCATCCAAATGAGACTTCTCTCCACCC 533
Qy 159 CysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGln 178
```



```
Db 181 TACCTTCGTCTGATGTTAGAAAGGAAAAACATTACACCCGAGAACAGCTTTTGATCATA 240
Qy 83 GluLeuHisAlaLysTyrGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
Db 241 GAACATTCATGCAAAAGTGGGGCAATAGGTGGTCCAAAATTGCAAAAGCATCTTCCAGGAAGA 300
Qy 103 ThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAla 122
Db 301 ACTGACATGAGATTAAAGACTTCTGGAGACAAAGGATCCAAAAGCACATTAAAGCAAGCT 360
Qy 123 GluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla----- 140
Db 361 GAGACTTCACAAACAT-----GGTAATTGAGAGAATAATGATCATCAAGCAAGCACT 414
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 415 AGTACTAGCAAAAGTGTCCACCATGGCACATCCAAATGAGACTTCTCTCCACCCCTCATAC 474
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 475 CAAGGAACCTTTGAGCCACTTCAACCTCAATTCCTTACAACTACT--GATCAATCAAGT 531
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp---Ser 199
Db 532 TGTGTACCACCACCAACGACAAACAACTATTGGAGCATCGAGGATATCTGTCGTCT 591
Qy 200 MetGlnLeuLeuAsnGlyAsp 206
Db 592 ATGCAATTACTCAATGGAGAT 612
```

RESULT 15

```
US-10-659-869-37
; Sequence 37, Application US/10659869
; Publication No. US20040040057A1
; GENERAL INFORMATION:
; APPLICANT: Fahg, Yiwen
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/659,869
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Glycine max
US-10-659-869-37
```

Alignment Scores:

Pred. No.:	9,57e-102	Length:	805
Score:	874.50	Matches:	168
Percent Similarity:	85.51%	Conservative:	9
Best Local Similarity:	81.16%	Mismatches:	23
Query Match:	76.85%	Indels:	7
DB:	18	Gaps:	5

US-10-659-869A-36 (1-206) x US-10-659-869-37 (1-805)

```
Qy 4 LysGlnGlnCys---LysThrSerGlnAspProGluValArgLysGlyProTrpThrMet 22
Db 1 AAAAAACCATGCAACTCATCATCTGATCTGAAGTGAGAAAGGACCATGGACCATG 60
Qy 23 GluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrpAsnSer 42
Db 61 GAAGAAGACTTGATCTTGATAAACTATATTGCAAAATCACGGTGAAGGTTGTTGGAACTCC 120
```

```
Qy 43 LeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsn 62
Db 121 TTAGCCAAAGCTTCTGCTCTCAACACGACGGAAGAGTTGTCCACTCCGTTGCTAAAC 180
Qy 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeuIleMet 82
Db 181 TACCTTCGTCTGATGTTAGAAAGAGAAACATTACACCCGAGGAACAGCTTTTGATCATA 240
Qy 83 GluLeuHisAlaLysTyrGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
Db 241 GAACATTCATGCAAAAGTGGGGCAATAGGTGGTCCAAAATTGCAAAAGCATCTTCCAGGAAGA 300
Qy 103 ThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAla 122
Db 301 ACTGACATGAGATTAAAGACTTCTGGAGAACAAAGGATCCAAAAGCACATTAAAGCAAGCT 360
Qy 123 GluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla----- 140
Db 361 GAGACTTCACAAACAT-----GGTAATTGAGAGAATAATGATCATCAAGCAAGCACT 414
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
Db 415 AGTACTAGCAAAAGTGTCCACCATGGCACATCCAAATGAGACTTCTCTCCACCCCTCATAC 474
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 475 CAAGGAACCTTTGAGCCACTTCAACCTCAATTCCTTACAACTACT--GATCAATCAAGT 531
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAspSerTrp---Ser 199
Db 532 TGTGTACCACCACCAACGACAAACAACTATTGGAGCATCGAGGATATCTGTCGTCT 591
Qy 200 MetGlnLeuLeuAsnGlyAsp 206
Db 592 ATGCAATTACTCAATGGAGAT 612
```

Search completed: September 28, 2005, 17:19:01

Job time : 641 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 28, 2005, 12:09:12 ; Search time 3166 Seconds
(without alignments)
2476.702 Million cell updates/sec

Title: US-10-659-869A-36

Perfect score: 1138
Sequence: 1 MDKXQCKTSQDPEVRKGPW.....NNINYSMEDSWQMLNGD 206

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10659869/runat.26092005.143445.11376/app.query.fasta_1.391
-DB=EST -QWMT=fascap -SURFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10659869 @CGN 1.1 3437 @runat.26092005.143445.11376 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	991	87.1	562	BM732121	BM732121 sal173b11
2	859.5	75.5	635	BQ146360	BQ146360 NF047E06F
3	851	74.8	501	BM527606	BM527606 sal63906
4	846	74.3	782	BF5658316	BF5658316 GM700005B
5	816	71.7	552	BM527774	BM527774 sal165903
6	813.5	71.5	560	BM732539	BM732539 sal178e08
7	808	71.0	669	EG457971	EG457971 NF037A10P
8	796	69.9	682	BQ146831	BQ146831 NF028G06F
9	795.5	69.9	568	BM527664	BM527664 sal64d10

10	784	68.9	523	1	AI930997	AI930997 sb45h07.Y
11	784	68.9	588	4	BI272897	BI272897 NF091A12F
12	779.5	68.5	544	4	BM527508	BM527508 sal162f08
13	774.5	68.1	542	4	BM528383	BM528383 sal57f09
14	751.5	66.0	659	7	CV130213	CV130213 Mdfert3088
15	749.5	65.9	638	7	CV466307	CV466307 Mdfert3092
16	726.5	63.8	805	7	CO113188	CO113188 GR_EB013
17	724	63.6	681	1	AI486576	AI486576 EST244897
18	721	63.4	761	1	AJ791605	AJ791605 AJ791605
19	715.5	62.9	522	5	BP032853	BP032853 BP032853
20	711	62.5	658	1	AJ790280	AJ790280 AJ790280
21	709.5	62.3	476	2	BE057370	BE057370 en02c04.Y
22	705	62.0	619	1	AJ554702	AJ554702 AJ554702
23	695.5	61.1	663	7	CK468561	CK468561 AMF121_B0
24	690.5	60.7	570	7	CV298677	CV298677 EST87136
25	673.5	59.2	592	2	AW928296	AW928296 EST307029
26	671	59.0	558	2	BE324639	BE324639 NF024C04P
27	670.5	58.9	470	2	AW156542	AW156542 ae28d11.Y
28	669	58.8	567	7	CO414782	CO414782 Mdfw2065k
29	667.5	58.7	642	4	BI978095	BI978095 pE09_Old
30	667	58.6	896	3	CNS0A1B5	EX829759 Arabidops
31	657	57.7	409	2	BE804790	BE804790 sb45f11.Y
32	656.5	57.7	616	7	CF349392	CF349392 fc3765.e
33	654.5	57.5	425	2	AW423958	AW423958 sh58e05.Y
34	642	56.4	503	5	BQ104458	BQ104458 g91307.e
35	639	56.2	500	1	AI897784	AI897784 EST267227
36	639	56.2	519	1	AI897681	AI897681 EST267124
37	637.5	56.0	502	5	BQ106505	BQ106505 fc0568.e
38	635	55.8	596	6	CB078113	CB078113 hJ63d11.g
39	627.5	55.1	447	4	BI273011	BI273011 NF097F02F
40	627.5	55.1	452	7	CN495741	CN495741 Mdfw2020d
c	625	54.9	803	4	BJ568794	BJ568794 BJ568794
41	621	54.6	395	2	BE805071	BE805071 sb37a07.Y
43	619	54.4	557	4	BJ553935	BJ553935 BJ553935
44	614	54.0	450	2	BE058947	BE058947 en23b01.Y
45	612	53.8	537	5	BQ584246	BQ584246 E011860-0

ALIGNMENTS

RESULT 1
BM732121
LOCUS
DEFINITION sal173b11.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl061-4678 5', similar to TR:Q39028 Q39028 ATMYB2.1, mRNA
sequence.
BM732121
BM732121.1 GI:19053454
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 562)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,B., Kohn,S., Shin,T., Jackson,I., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES

```

source
1..562
  /organism="Glycine max"
  /mol_type="mRNA"
  /cultivar="Raiden"
  /db_xref="taxon:3847"
  /clone="SOYBEAN CLONE ID: Gm-c1061-4678"
  /tissue_type="mature flowers of field grown plants"
  /lab_host="DH10B"
  /clone_lib="Gm-c1061"
  /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      1.03e-99      Length:      562
Score:          991.00        Matches:    181
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      87.08%      Indels:      0
DB:               4           Gaps:         0

```

US-10-659-869A-36 (1-206) x BM732121 (1-562)

```

QY 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
DB 20 ATGGATATAAACAACAGTGTAAAGACGTCCTCAAGATCTCTGAAGTACGAAAGGGCCTTGG 79
QY 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
DB 80 ACAATGGGAAGACCTTGATCTTGATGAACATATATTGCAATCATCTGGGGAAGTCTTTGG 139
QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
DB 140 AACTCTTTGGCCAAAGCTGCTGGTCTCAAAACGTAACGGAAGAGTTGCCGGCTAAGGTGG 199
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluLeuLeu 80
DB 200 CTAAATTTACCTCCGTCCTGATGTTAGAGAGGGAATATTACACCGAGAACACTTTTG 259
QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
DB 260 ATTATGGAGCTCCACGCAAGTGGGAAACAGTGTGTCARAAATTTGCCAAGCATCTACCT 319
QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
DB 320 GGAAGGACTCTGATAATGAGATCAAGAACTATTTCGAGGACAAGGATCCGAAGCACATCAAG 379
QY 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
DB 380 CAAGCTGAGAACTTTTCAGCAACAGAGTAGTAATAATTCGAGATAAATGATCACCAGCT 439
QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
DB 440 AGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGGAGATGATTCTCCACCTGTTAT 499
QY 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
DB 500 CAAGGAATGTTAGAGCCATTTTCAACTCAGTTTCCCTACAAATTAATCTCCTGATCAATCCAGT 559

```

```

QY 181 Cys 181
DB 560 TGT 562

```

RESULT 2

```

BQ146360
LOCUS      BQ146360
DEFINITION NF047806FL1F1050 Developing flower Medicago truncatula cDNA clone
EST.
ACCESSION BQ146360
VERSION   BQ146360.1 GI:20283419
KEYWORDS  EST.
SOURCE    Medicago truncatula (barrel medic)
ORGANISM  Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 635)
AUTHORS   Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE      Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
JOURNAL    Unpublished (2001)
COMMENT    Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 635 Std Error: 0.00
Plate: 047 Row: E Column: 06
Seq primer: TCACACGAGAAACAGCTATGAC.

```

FEATURES

```

source
1..635
  /organism="Medicago truncatula"
  /mol_type="mRNA"
  /db_xref="taxon:3880"
  /clone="NF047806FL"
  /tissue_type="Developing flowers"
  /dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
  /clone_lib="Developing flower"
  /note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF+ (Stratagene). Excised plasmids were plated
using SOLR cells."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      5.27e-85      Length:      635
Score:          859.50        Matches:    167
Percent Similarity: 85.29%    Conservative: 7
Best Local Similarity: 81.86% Mismatches: 18
Query Match:     75.53%      Indels:      12
DB:              5           Gaps:         5
US-10-659-869A-36 (1-206) x BQ146360 (1-635)
QY 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
DB 37 ATGGATAAAAAA---CCATGCAATCATCTCAAGATCTCTGAAGTACGAAAGGGCCATGG 93

```

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40
 Db 94 ACCATGGAAGAAGACTTGATTTAATCAATATATATGCAATCATGTGAGGTGTTTGG 153
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
 Db 154 AATTCCTTAGCCAAAGCTGCTGCTTAACGTCACAGGAAAAGTTCCAGGCTTCGATGG 213
 Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 214 TTAACACTACCTTCCTGCAGATGTTAGAGAGGGAATATATACACCTGAGGAACAACCTTTTG 273
 Qy 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
 Db 274 ATCATGGAACCTTCATGCAAAAGTGGGAAATAGGTGGTCCAAAATTCGAAAGCATCTTCCA 333
 Qy 101 GlyArgThrAspAsnGluLeuLysAsnTyrTrrpArgThrArgIleGlnLysHisIleLys 120
 Db 334 GGAAGAAGCTGACATGAGATTAGAAATTTTGGAGGACTAGGATACAAAGACATTTAAG 393
 Qy 121 GlnAlaGlu-----AsnPheGlnGlnSerSerAsnAsnSerGluIle 135
 Db 394 CAAGTTGATCATCTCCACCAAAATAATTTCCACCAAAATGAGT-----TTAGAGATA 444
 Qy 136 AsnAsp---HisGlnAlaSerThrSer-HisValSerThrMetAlaGluProMetGluMe 154
 Db 445 AATGATCATCATCAATCAAGTAGTAGGCCCAAGTTTCCAAACATTTAGAGCCAAATGGAATC 504
 Qy 154 tTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIle 174
 Db 505 TTACTCTCCACCTTCATACCAGGAACCTTTGGAGCCATTTCCAGCTCAATTTCCAAACAT 564
 Qy 174 eAsnProAspGln---SerSerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTrrpSe 193
 Db 565 CAATGATCATCATCCCTCAAGTTGTTGACAAATGACAAATACAAATATTATTCGAG 624
 Qy 193 rMetGluAsp 196
 Db 625 CATGGAGAT 634

BM527606 501 bp mRNA linear EST 06-JUL-2004
 sal6306.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-c1061-3780 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA
 sequence.

BM527606
 BM527606.1 GI:18733434
 EST.
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Fabaceae;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 501)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
 Seq primer: -4ORP from Gibco
 High quality sequence stop: 421.

FEATURES

source

1..501
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Raiden"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-c1061-3780"
 /tissue_type="mature flowers of field grown plants"
 /lab_host="DH10B"
 /clone_11b="Gm-c1061"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from mature flowers of field grown plants for the cultivar
 Raiden. Complementary DNA was synthesized from mRNA using
 a primer consisting of a poly(dT) sequence with a XhoI
 restriction site. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the pBluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."

ORIGIN

Alignment Scores:
 Pred. No.: 3 31e-84 Length: 501
 Score: 851.00 Matches: 157
 Percent Similarity: 96.34% Conservative: 1
 Best Local Similarity: 95.73% Mismatches: 6
 Query Match: 74.78% Indels: 0
 DB: 4 Gaps: 0

US-10-659-869A-36 (1-206) x BM527606 (1-501)

Qy 1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
 Db 10 ATGGATAAAAAACAACATGTGCAACAGCTCTCAAGATCTCTGAAGTGAGAAAAAGGACCTTG 69
 Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40
 Db 70 ACGATGGAAGAAGACTTGATCTTGTATCAACTATATTGCAAAATCATGGGAAGGTGTTTGG 129
 Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
 Db 130 AATTCCTTAGCCAAAGCTGCTGCTCAACAGTACCGGAAGAGTTCGCCGCTAAGGTGG 189
 Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
 Db 190 CTAAACTACCTCCGTCCTGATGTTAGAGAGGGAATATTACACCCGAGGAACAACCTTTTG 249
 Qy 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
 Db 250 ATCATGAGGCTTCACGCAAAAGTGGGAAACAGGTGGTCCAAAATTCGCAAGCATCTACCT 309
 Qy 101 GlyArgThrAspAsnGluLeuLysAsnTyrTrrpArgThrArgIleGlnLysHisIleLys 120
 Db 310 GGTAGGACAGATATATGATCAAGAATATTGGAGGACCAAGGATCCAGAGACATCAAG 369
 Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
 Db 370 CAAGCTGAGAACTTTCAGCAACAAATTTAGTAATAACTCTGAGATAAATGATCACCAGCT 429
 Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
 Db 430 AGCACTAGCCATGTTTCTTACCATGGCTGAACCCCATGGAGACCTATTCTCCACCCTTTAT 489
 Qy 161 GlnGlyMetLeu 164
 Db 490 CAAGGAATGTTA 501

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 426.

FEATURES

source
1. .552
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Raiden"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-4134"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Alignment Scores:
Pred. No.: 2,98e-80 Length: 552
Score: 816.00 Matches: 152
Percent Similarity: 95.62% Conservativeness: 1
Best Local Similarity: 95.00% Mismatches: 5
Query Match: 71.70% Indels: 2
DB: 4 Gaps: 2

US-10-659-869A-36 (1-206) x BM527774 (1-552)

Qy 49 LeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspVal 68
Db 1 CTCAAAGCTACCGAAGAGTGTCCGGCTAAGTGGCTAACTACCTCCGCTCGATGTT 60
Qy 69 ArgArgGlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrp 88
Db 61 AGAAGAGGGAATATTACACCGGAGAACACTTTTGTATCATGGAGCTTCACGCAAAAGTG 120
Qy 89 GlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluLys 108
Db 121 GGAAACAGGTGGTCCAAAATTTGCCAACATCTACCTGTAGGACAGATAATGATCAAG 180
Qy 109 AsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGln 128
Db 181 AACTATTGGAGGACGAGATCCAGACACATCAAGCAAGCTGAGAACTTTTCACCAACA 240
Qy 129 SerSerAsnAsnSerGluLeuAsnAspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 241 ATTAGTAATAACTCTGAGATAAATGATCACCAGCTAGCCTAGCCATGTTCTTACCATG 300
Qy 149 AlaGluProMetGluMetTyrSerProProCysTyrTrpGlnGlyMetLeuGluProPheSer 168
Db 301 GCTGAACCCATGGAGACCTATCTCCACCTTTTATCAAGGAATGTTAGAGCCATTTC 360
Qy 169 Thr--GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsn 187
Db 361 TCAATTCAGTCCCAACAAATTAATCTTGATCAATCCAGTGTGTGTACATGACACAC 420

Qy 188 Asn--IleAsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db 421 AACAGCATTAACTATTGGAGCATGGAGGATATCTGGTCAATGCAGTTACTGAACGGGAT 480

RESULT 6

BM732539
LOCUS
DEFINITION
Gm-cl061-5223 5' similar to TR:Q3028 Q3028 ATMYB2. i, mRNA sequence.

ACCESSION

BM732539.1 GI:19053872

VERSION

BM732539

KEYWORDS

Glycine max (soybean)

SOURCE

Glycine max

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE

1 (bases 1 to 560)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 430.

Location/Qualifiers

1. .560

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Raiden"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl061-5223"

/tissue_type="mature flowers of field grown plants"

/lab_host="DH10B"

/clone_lib="Gm-cl061"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

Alignment Scores:

Pred. No.: 5.78e-80 Length: 560

Score: 813.50 Matches: 154

Percent Similarity: 85.71% Conservativeness: 8

Best Local Similarity: 81.48% Mismatches: 22

Query Match: 71.49% Indels: 5

DB: 4 Gaps: 3

ORIGIN

US-10-659-869A-36 (1-206) x BM732539 (1-560)

```

Qy 9 ThrSerGlnAspProGluValArgLysGlyProTrrpThrMetGluGluAspIleuIleu 28
Db 3 TCATCTCATGATCCTGAAGTGAAGAGGACCATGGACCATGGGAAGACATTCATCTTG 62
Qy 29 MetAenTyrIleAlaAsnHisGlyGluValTrrpAsnSerLeuAlaLysAlaAlaGly 48
Db 63 ATAAATATATGTCAAATTCACGGTGAAGGTGTGGAACTCTTAGCCAAAGCTTCTGGT 122
Qy 49 LeuLysArgAsnGlyLysSerCysArgLeuArgTrrpLeuAenTyrLeuArgProAspVal 68
Db 123 CTCAAACGAACGGGAAGAGTGTGCATCCGTTGGCTAAAGTAACTACCTTCGCTCATGTT 182
Qy 69 ArgArgGlyAsnIleThrProGluGlnLeuIleMetGluLeuHisAlaLysTrp 88
Db 183 AGAGAGGAATTTTACACCCGAGGAACAGCTTTTGATCATAGAACTTCATGCAAGTGG 242
Qy 89 GlyAenArgTrrpSerLysIleAlaLysHisLeuProGlyArgTrrpAspAsnGluLys 108
Db 243 GGCAATAGTGGTCCAAATTCGAAGATCTTCCAGGAAGACTGACCAATGAGATTAG 302
Qy 109 AenTyrTrrpArgThrArgIleGlnLysHisLysGlnAlaGluAsnPheGlnGln 128
Db 303 AACTTCTCGAGAACCAAGGATCCAAAGCACATTAAGCAAGCTGAGACTTCACAACAACAT 362
Qy 129 SerSerAsnAsnSerGluIleAsnAspHisGlnAla-----SerThrSerHisValSer 146
Db 363 -----GGTAATTCAGAAATAATGATCATCAAGCAAGCATTAGTACTAGCAAGTGTCC 416
Qy 147 ThrMetAlaGluProMetGluMetTrrpSerProProCysTrrpGlnGlyMetLeuGluPro 166
Db 417 ACCATGGCACATCCAAATGAGACTTCTCTCCACCTCATCAAGGAACTTTTGAGCCA 476
Qy 167 PheSerThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsn 186
Db 477 TTCCAACCTCAATCCCTACAACTACT---GATCAATCAAGTTGTTGTACCAACCAAC 533
Qy 187 AsnAsnIleAenTyrTrrpSerMetGlu 195
Db 534 GACAACAACAATATTGGAGCATCGAG 560

```

RESULT 7
 BG457971
 LOCUS
 DEFINITION
 NP037A10P1F1070 Phosphate starved leaf Medicago truncatula cDNA
 clone NF037A10PL 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Medicago truncatula (barrel medic)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; euroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 669)
 Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula phosphate-starved leaf library
 Unpublished (2000)

Contact: Harrison MJ
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7325
 Fax: 580 221 7380
 Email: mjharrison@noble.org
 Insert Length: 669 Std Error: 0.00
 Plate: 037 row: A column: 10
 Seq primer: TCACACGAGAACAGCTATGAC.
 Location/Qualifiers

FEATURES

source

1. .669
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF037A10PL"
 /tissue_type="leaf"
 /dev_stage="trifoliolate"
 /clone_lib="Phosphate starved leaf"
 /note="Vector: Lambda zap; At the trifoliolate stage, M.
 truncatula plants were transplanted to phosphate-free sand
 and grown for a further 30 days. During this 30 day
 period, the plants were fertilized twice weekly with 1/2
 Hoglands solution containing only 20uM potassium
 phosphate. RNA was prepared from above ground tissues."

ORIGIN

Alignment Scores:	3.06e-79	Length:	669
Pred. No.:	808.00	Matches:	159
Score:	81.86%	Conservative:	17
Percent Similarity:	73.95%	Mismatches:	19
Best Local Similarity:	71.00%	Indels:	20
Query Match:	4	Gaps:	7

US-10-659-869A-36 (1-206) x BG457971 (1-669)

Qy	1	MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp	20
Db	50	ATGGACAAAAA---GAGTGCAGTTCTTCAAGATCTCTGATGTACGAAGGGACCATGG	106
Qy	21	ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp	40
Db	107	ACATGGAGAGACTTGAATCTTGATCAATATATGCAATCATGTGGAAGGTGTTGG	166
Qy	41	AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp	60
Db	167	AACTCCTTGGTAAATCTGCTGCTTAAACGTACCGGAAAGAGTTGCAGGCTTCGGTGG	226
Qy	61	LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu	80
Db	227	CTAAACTATCTTCACCGGATGTAGACGAGGAATATTACCTTGAGGAACAATCTCTTG	286
Qy	81	IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro	100
Db	287	ATCATTTGAATCTCATGCTAGTGGGAAACAGGTGTCGAAATTTGCGAAACATCTACCG	346
Qy	101	GlyArgThrAspAsnGluIleLysAsnTyrTrrpArgThrArgIleGlnLysHisLys	120
Db	347	GGAAAGACCGATAATGAGATAAGAACTACTCTGGAGGACAAGGATCCAAAAGCATATCAAG	406
Qy	121	GlnAlaGluAsnPheGln---GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln	139
Db	407	CAAGCTGAGAACTCTCAATCTCAAGACAAGGTTCATATTTCAATTAATTAATGATGATAAT	466
Qy	140	-----AlaSerThrSerHisValSerThrMetAlaGluProMetGlu	153
Db	467	AATAATAATCAACTTATAGGAAGCACAGCCAAATTTCCAACATTTGCTGAACCTATGGAC	526
Qy	154	-----MetTyrSerProProCysTrrpGlnGlyMetLeuGluProPheSerThr	169
Db	527	ACAAATATTATATCACCACCTCTTACCAAGGAATTTTGGAAACCATTTTCCACCTCATCAT	586
Qy	170	GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnIle	189
Db	587	CAGTTCCCTACAACTCTCT---GATCAATCAGGTGTCTCT-----	622
Qy	190	AsnTyrTrrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsn	204
Db	623	AACTACTGGACATGGAAGATCTCTGTGCTCACTTCAACTACTATAAT	667

RESULT 8
 BQ146831
 LOCUS
 BQ146831
 682 bp
 mRNA
 linear
 EST 24-APR-2002

```

DEFINITION  NF028G06FL1F1051 Developing flower Medicago truncatula cDNA clone
ACCESSION   NF028G06FL 5', mRNA sequence.
VERSION     BQ146831
KEYWORDS    BQ146831.1 GI:20283890
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 682)
AUTHORS    Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
            Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.
TITLE       Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula flower library
JOURNAL     Unpublished (2001)
COMMENT     Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 224 6650
            Fax: 580 224 6692
            Email: gdmay@noble.org
            Insert length: 682 Std Error: 0.00
            Plate: 028 row: G column: 06
            Seq primer: TCACACGAGAAACAGCTATGAC.

FEATURES             Location/Qualifiers
     source           1..682
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /db_xref="taxon:3880"
                     /clone="NF028G06FL"
                     /tissue_type="Developing flowers"
                     /dev_stage="Developmentally pooled. Contains a mixture of
                     very young, developing, fully-opened flowers and flowers
                     in early transition into pods."
                     /clone_lib="Developing flower"
                     /note="Vector: Lambda Zap; cDNA was prepared from polyA+
                     enriched, pooled samples of equivalent amounts of total
                     RNA from very young, developing, fully-opened flowers and
                     flowers transitioning into pods. The cDNA was
                     directionally ligated into the Uni-Zap XR vector
                     (Stratagene) and packaged using the Gigapack III Gold
                     packaging extracts. Phagemids containing cDNA inserts were
                     in vivo excised from the recombinant Uni-Zap XR vector
                     using ExAssist helper phage and the E. coli strain
                     XL1-Blue MRF' (Stratagene). Excised plasmids were plated
                     using SOLR cells."

ORIGIN
Alignment Scores:
Pred. No.:      6.8e-78      Length:      682
Score:          796.00      Matches:    160
Percent Similarity: 81.19%      Conservative: 17
Best Local Similarity: 73.39%      Mismatches: 20
Query Match:    69.95%      Indels:    21
DB:             5          Gaps:      7

US-10-659-869A-36 (1-206) x BQ146831 (1-682)

Qy      1  MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
Db      47  ATGCACAAAAA---GAGTGCAGTCTTCACAGATCTGTGATGACGAAAGGACCATGG 103

Qy      21  ThrMetGluGluAspLeuLeuMetAsnTyrTrileAlaAsnHisGlyGluGlyValTyr 40
Db      104 ACAATGGAAGAAGACTTGATCTTGATCAACTATATATGCAATCATGTGTGAAGGTGTTGG 163

Qy      41  AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
Db      164 AACTCCTTGGCTAAATCTGTGGTCTTAAACGTAACCGAAGAGATTGACGCTTCGGTGG 223

```

```

Qy      61  LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db      224 CTAAACTATCTTCGACCGGATGTTAGACGAGGGAATATTACACCTGAGGAACAACTCTTG 283

Qy      81  IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro 100
Db      284 ATCAATGAACTTCACTGTAAGTGGGAACAGGTGTGCGAAATTTGCGAAACATCTACCG 343

Qy      101 GlyArgThrAspAsnGluLysAsnTyrTyrArgThrArgIleGlnLysHisIleLys 120
Db      344 GGAAGAACCGATAATCAGATAAAGAACTACTCGAGGACAAAGGATCCAAAAGCATATCAAG 403

Qy      121 GlnAlaGluAsnPhenGln---GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139
Db      404 CAAAGCTGAGAACTCTCAATCTCAAGAACCAAGGTTCTGATATTCAATATAAATGATGATAAT 463

Qy      140 -----AlaSerThrSerHisValSerThrMetAlaGluProMetG1 153
Db      464 AATAATAACTAATTATAGGAAGCACACAGCCAAATTTCCAACTATGCTGAACCTATGGA 523

Qy      153 u-----MetTyrSerProCysTyrGlnGlyMetLeuGluProPhaSerThr----- 169
Db      524 CACAATAATATTATATACACCCTCTTTACCAAGGAATTTTGGAAACCAATTTCCACCTCATCA 583

Qy      170 -GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnI1 189
Db      584 TCAGTTCCCTTACAACTCTCT---GATCAATCA-----GGTTGNTC 619

Qy      189 eAsnTyrTyrSerMetGluAspSerTyrSerMetGlnLeuLeuAsnGlyAsp 206
Db      620 TAACTACTGGAGCATGGAAGATCTCTTGTCACCTCACTAACTAAATGGNGAT 671

RESULT 9
BM527664
LOCUS      BM527664
DEFINITION 568 bp mRNA linear EST 06-JUL-2004
            sal64dl0.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1061-4003 5' similar to TR:Q39028 Q39028 ATMB2. ; mRNA
            sequence.
ACCESSION  BM527664.1 GI:18733532
VERSION    BM527664
KEYWORDS   EST.
SOURCE     Glycine max (soybean)
ORGANISM   Glycine max
            Glycine max
REFERENCE   1 (bases 1 to 568)
AUTHORS    Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,
            Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
            Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
            Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
            Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
            McCann, R., Waterston, R. and Wilson, R.
            Public Soybean EST Project
            Unpublished (1999)
            Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            When it has been determined, an EST from the other end of this
            clone is listed in the 'Other ESTs on clone' field. This clone is
            available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
            57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source           1..568
                     /organism="Glycine max"
                     /mol_type="mRNA"

```

/cultivar="Raiden"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-4003"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Alignment Scores:
Pred. No.: 5, 94e-78 Length: 568
Score: 795.50 Matches: 153
Percent Similarity: 87.91% Conservative: 7
Best Local Similarity: 84.07% Mismatches: 19
Query Match: 69.90% Indels: 3
DB: 4 Gaps: 3

US-10-659-869A-36 (1-206) x BM527664 (1-568)

Qy 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 29 ATGGACAAGAGCTTGGC---AACACGCTCTCATGATCTCTGAAGTCAGAAAGGGGCGCATGG 85
Qy 21 ThrMetGluGluAspLeuMetAsnTyrIleAlaAsnHisGluGluValTrp 40
Db 86 ACAATGGGAAGAGACTTAATCTTGTATACCTATATTTGCCAATCATCGGGGAAGGGGTTTGG 145
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 146 AACTCTTTGGCCAGCGCTGCTGGACTTAAACGTACCGGAAGAGATGGCGGCTCGGTGG 205
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 206 CTAACACTACCTCCGTCCTGATGTTAGAAGAGGGAATATTACACCGAGGAACAGCTTTTG 265
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 266 ATCATGGGAACCTTCATGCAAGTGGGGAACAGGTGTCCTCAAAATTTGCCAAGCATCTACCC 325
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 326 GGAAGAGCTGATTAATGAGATTAGAACTACTGGAGAGACAGGATTCAGAGCAGCTCAAG 385
Qy 121 GlnAla---GluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139
Db 386 CAAGCTTCAGCAGCTTCAGCAACAGAGTAGTAATTCGTAGATAATTTATCATCCCCAA 445
Qy 140 AlaSerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCys 159
Db 446 GCTTCACCTAGCCAAGTGTCCACCATGGCGCCCATAGAACCTATTCTCCACCCAGT 505
Qy 160 TyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 179
Db 506 TATCAAGGAGATGTTAGATCCATTTCAATTCAATTGATTTCCCAACA---AATCCTCATCTCT 562
Qy 180 SerCys 181
Db 563 ACTTGT 568

RESULT 10

AI930997
LOCUS 523 bp mRNA linear EST 24-JUL-2004

DEFINITION sb45h07.y1 Gm-cl015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl015-278 5' similar to TR:Q39028 Q39028 ATMYB2. i, mRNA sequence.
ACCESSION AI930997
VERSION AI930997.1 GI:5666961
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
REFERENCE 1 (bases 1 to 523)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Other ESTs: BE658316 corresponding to Gm-ri070-1759 (3')
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 422.
Location/Qualifiers
1. 523
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-278"
/tissue_type="Mature flowers, field grown plants"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl015"
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

ORIGIN

Alignment Scores:
Pred. No.: 1e-76 Length: 523
Score: 784.00 Matches: 146
Percent Similarity: 94.84% Conservative: 1
Best Local Similarity: 94.19% Mismatches: 8
Query Match: 68.89% Indels: 0
DB: 1 Gaps: 0

US-10-659-869A-36 (1-206) x AI930997 (1-523)

Qy

1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
|||||

Db 58 ATGATATAAAACAACTGTGCAACACGCTCTCAAGATCTCTGAAGTGAGAAAAGGACCTTGG 117

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40

Db 118 ACGATGGAATAAGACTTTCATCTTGATCAACTATATTGCAATATCTGGGAAGGTGTTGG 177

Qy 41 AsnSerIleuAlaLysAlaAlaGlyLeuLysArgGlnGlySerCysArgLeuArgTrp 60

Db 178 AATTCTTTGGCCAAAGCTGCTGCTCAACAGCTACCGAAAGAGTTCGCCGCTAAGGTGG 237

Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80

Db 238 CTAACTACCTCCGCTCGCTGATGTAGAGAGGAGATATTACCCCGAGAACAACTTTTG 297

Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerIleAlaLysHisLeuPro 100

Db 298 ATCATGGAGCTTCACGCCAAGTGGGAAACAGGTGCTCAAAATTTGCCAAGCATCTACCT 357

Qy 101 GlyArgThrAspAsnGluLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120

Db 358 GGTAGGACAGATAATGAGATCAAGAACTATTGGAGGACCCAGGATCCAGAACATCAAG 417

Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140

Db 418 CAAGCTGAGAACTTTCAGCAACAATCAGCAATTAATCTGAGATTAATGATCACAAGCT 477

Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyr 155

Db 478 AGCACTAGCCATGTTTCTACCATGGCTGACCCCATGGAGACCTAT 522

RESULT 11

BI272897

LOCUS

DEFINITION

NP091A12FL1088 Developing flower Medicago truncatula cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 688)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula flower library

Unpublished (2001)

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 224 6650

Fax: 580 224 6692

Email: gdmay@noble.org

Insert Length: 688 Std Error: 0.00

Plate: 091 Row: A Column: 12

Seq primer: TCACACGGAACACGCTATGAC.

FEATURES

Location/Qualifiers

1..688

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NP091A12FL"

/tissue_type="Developing flowers"

/dev_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."

/clone_lib="Developing flower"

/note="Vector: Lambda Zap; cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total

RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XLI-Blue MRF⁺ (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Alignment Scores: 1.49e-76 Length: 688

Pred. No.: 784.00 Matches: 158

Score: 81.57% Conservative: 19

Percent Similarity: 72.81% Mismatches: 20

Best Local Similarity: 68.89% Indels: 21

Query Match: 4

Gaps: 7

US-10-659-869A-36 (1-206) x BI272897 (1-688)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20

Db 49 ATGGACAAAAA---GAGTGCAGTTCCTTCACAGATCTCTGATGTCGAAAGGACCATGG 105

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40

Db 106 ACAATGGAGAAGACTTGTATCTTGATCAATATATTGCCAATCATGGAAGAGGTGTTGG 165

Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60

Db 166 AACTCTCTGGCTAAATCTGCTGTCTTAAACGTCACCGAAAGAGTTCCAGGCTTCGGTGG 225

Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80

Db 226 CTAAACTATCTTCGACCGGATGTTAGACGAGGATATATTACCTCTGAGGAACAACACTCTTG 285

Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100

Db 286 ATCATTTGAATTCATGCTAAGTGGGAAACAGGTGGTCGAAATTTGCGAAACATCTACCG 345

Qy 101 GlyArgThrAspAsnGluLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120

Db 346 GGAAGAACCCGATATAGATATAAGAACTACTTGGAGGACAAAGGATCCAAAAGCATATCAAG 405

Qy 121 GlnAlaGluAsnPheGln---GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139

Db 406 CAAGCTGAGAACTCTCAATCTCAAGAACAAAGGTCTCTGATATTCAATTAATGATGATAAT 465

Qy 140 -----AlaSerThrSerHisValSerThrMetAlaGluProMetGlu 153

Db 466 AATAATAATCAACTTATAGGAAGCACAAAGCCAAATTTCCAAACATTTGCTGAACCTATGGAC 525

Qy 154 -----MetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThr----- 169

Db 526 ACAAAATATTATATACCAACCCCTCTTACCAAGGAATTTTGGAAACCATTTCCACCTCATCAT 585

Qy 170 GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnAsnIle 189

Db 586 CAGTTCCCTCAATCTCT---GATCAATCAGTTGTTCT----- 621

Qy 190 AsnTyrTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206

Db 622 AACTACTGGACCATGGAAAATCTCTTGTCACTTCA-CTACTAAATGGNGAT 671

RESULT 12

BM527508

LOCUS

DEFINITION

sal62f08.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1061-3735 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA sequence.

ACCSSION

VERSION

BM527508

BM527508.1 GI:18733261

	EST.	Glycine max (soybean)	
KEYWORDS	SOURCE	Glycine max	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
REFERENCE	AUTHORS	1 (bases 1 to 544) Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson.R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Seq primer: -4ORP from Gibco High quality sequence stop: 422. Location/Qualifiers 1..544 /organism="Glycine max" /mol_type="mRNA" /cultivar="Raiden" /db_xref="taxon:3847" /clone="SOYBEAN CLONE ID: Gm-cl061-3735" /tissue.type="mature flowers of field grown plants" /lab host="DH10B" /clone.lib="Gm-cl061" /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."	
FEATURES	source	Pred. No.: 3.37e-76 Length: 544 Score: 779.50 Matches: 150 Percent Similarity: 88.70% Conservative: 7 Best Local Similarity: 84.75% Mismatches: 17 Query Match: 68.50% Indels: 3 DB: 4 Gaps: 3	
ORIGIN		US-10-659-869A-36 (1-206) x BM527508 (1-544)	
		Alignment Scores:	
		QY 1 MetAspLysLeuGlncCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20	
Db		14 ATGGACAGAAAGCTTGCC---AACACGTCTCATGTCCTGAATCGAGAAGGGGCCATGG 70	
QY		21 ThrMetGluGluAspLeuLeuMetAsnTyrrileAlaAenHisGlyGluGlyValTrp 40	
Db		71 ACAATGGAAGAGACTTAATCTTTGATCACCTATTATGCATATCAAGCGGGAAGGGGTTTTG 130	
QY		41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAenGlyLysserCyserAtgLeuArgTrp 60	

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Alignment Scores:
 Pred. No.: 1,21e-75 Length: 542
 Score: 774.50 Matches: 150
 Percent Similarity: 85.71% Conservative: 6
 Best Local Similarity: 82.42% Mismatches: 21
 Query Match: 68.06% Indels: 5
 DB: 4 Gaps: 3

US-10-659-869A-36 (1-206) x BMS28383 (1-542)

```

Qy 1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
Db 6 ATGACAAAACCAACATCATCTCTCATGCTCTGAAGTGAGAAAGGGACCATGG 65

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40
Db 66 ACCATGGAAGAAGACTTGTATCTGATAAATATATATGCAATCATCGGTGAAGTGTGG 125

Qy 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
Db 126 AACTCTTAGCAAGACTTCTGGTCTCAACGACGGAAGAGTGTGCGACTCCGTGG 185

Qy 61 LeuAsnTyrLeuArgProAspValArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 186 CTAAATACCTTCCTCTGATGTTAGAGAGGAAACATTCACCCGAGGACACGCTTTTG 245

Qy 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
Db 246 ATCATAGAACTTCATGCAAGTGGGCAATAGTGTGCTCAAAATTCGAAAGCATCTTCCA 305

Qy 101 GlyArgThrAspAsnGluLeuLysAsnTyrTrrpArgThrArgIleGlnLysHisIleLys 120
Db 306 GGAAGAACTGCAATGAGATTAGAACTTCTGGAGAACAGGATCCAAAGACATTAAG 365

Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 366 CAAGCTGAGACTTCACACACACAT-----CGTAATTCAGAGATATATGATCATCAGCA 419

Qy 141 -----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerPro 158
Db 420 AGCACTAGTACTAGCAAGTGTCCACCATGTCACATCCAAATGAGACTTCTCTCCACCC 479

Qy 159 CysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGln 178
Db 480 TCATACCAAGGAACCTTTTGAGCCATTCCAACTCAATTCCTTCAATCACT---GATCAA 536

Qy 179 SerSer 180
Db 537 TCAGT 542

```

RESULT 14
 CVI30213
 LOCUS
 DEFINITION Mdftr30881.07.y1 Mdftr Malus x domestica cDNA clone Mdftr30881.07.5, similar to TR:Q39028 Q39028 ATMYB2.1; mRNA sequence.
 CVI30213
 ACCESSION CVI30213.1 GI:51865733
 VERSION
 KEYWORDS EST.
 SOURCE Malus x domestica (cultivated apple)

ORGANISM

Malus x domestica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids; eustosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE

AUTHORS

1 (bases 1 to 659)
 Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A., Aldwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Orvis, K., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I., Teagareishvili, R., Kennedy, S., Waterston, R. and Wilson, R.
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)

TITLE

JOURNAL

COMMENT

Contact: Schuyler S. Korban
 Apple Functional Genomics Grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library materials provided by: Schuyler S. Korban Library
 constructed by: K. Gasic Library sequenced by: Washington
 University Genome Sequencing Center
 WashU EST name: aal19e04.y1
 Seq primer: -40UP from Gibco.

FEATURES

source

1..659
 Location/Qualifiers
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /db_xrefs="taxon:3750"
 /clone="Mdftr3088107"
 /lab_host="DH10B ampicillin resistant"
 /clone_lib="Mdftr"
 /notes="Vector: pBluescript II SK (+); Site 1: NotI; Site 2: EcoRI; Total RNA was extracted separately from each stage [young fruitlet (<1cm), young fruitlet (1 cm dia.), young fruitlet (12cm dia.), maturing fruit I, maturing fruit II, mature fruit], using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo18(dT) primer with an identifying tag sequence (see table below). cDNA's from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (young fruitlet) insert 18(A)TCGTG; Stage 2 (young fruitlet 1cm dia) insert 18(A)TCGTG; Stage 3 (maturing fruit I) insert 18(A)TCGA; Stage 4 (maturing fruit II) insert 18(A)TCGA; Stage 5 (maturing fruit I) insert 18(A)TCGA; Stage 6 (mature fruit) insert 18(A)TCGTG; Tag identification when sequencing from 3' end: Stage 1 (young fruitlet) CAGCA18(T) insert; Stage 2 (young fruitlet 1cm dia) CAGCA18(T) insert; Stage 3 (young fruitlet 12cm dia) ACCGA18(T) insert; Stage 4 (maturing fruit I) TCGCA18(T) insert; Stage 5 (maturing fruit II) TCGCA18(T) insert; Stage 6 (mature fruit) ACGCA18(T) insert. Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adapters at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene). Identification of adaptors and tags in 5'-end sequenced clones:
 <Vector>...TAAGCTT<End Vector><Start
 EcoRI adaptor>GATATCGAATTCATGTTGTTGG <End
 EcoRI adaptor><Start Insert>...AAAAAAAAAAAAAAAAAAAA-End
 Insert> <Start tag>TGCAG<End Tag><Start
 NotI site>Vector>GGCGCCGACCCCGG... The total number of white colony forming units (cfu) in the primary library before amplification was 2.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 4 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming

sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 4 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 5.6x10⁶ cfu. Background of empty clones was less than 1%.

ORIGIN

Alignment Scores:
 Pred. No.: 5,79e-73 Length: 659
 Score: 751.50 Matches: 149
 Percent Similarity: 77.51% Conservative: 13
 Best Local Similarity: 71.29% Mismatches: 30
 Query Match: 66.04% Indels: 17
 DB: 7 Gaps: 5

US-10-659-869A-36 (1-206) x CV130213 (1-659)

```

Qy 1 MetAspLysLysGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db ATGGATAAAATPCCATTCAATAGTTCATCCAGGATGTTGAAGTCAGAAAAGGCCATGG 83
Qy 21 ThrMetGluGluAspLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrp 40
Db ACCATGGAGAAGATTGATTCATCAACTATATGCAATATGCAATCATGGTGAAAGGTATGG 143
Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
Db 144 AACTCCCTAGCCAAAGCTGCTGTCTCAACGCTACTGGGAAGAGCTGCCACTCCGGTGG 203
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluClnLeuLeu 80
Db 204 CTGAATTATCTCGCGCTGATGTTCCGAGAGGAAATATCACTCTCTGAGGAACAACATTTG 263
Qy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 264 ATTATGGAACTGCATGCAAGTGGGGTACAGGTGTCGAAATCGAAACATCTACCG 323
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisIleLys 120
Db 324 GGAAGGACTGATAACGAAATCAAGAACTATTGGAGGACTAGAAATTCAAAGCACATTAA 383
Qy 121 GlnAlaGluAsnPheClnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 384 CAAGCTGAGAATATTACACGAGCAAAAGC-----TCTGAGGTCAATGAT---CAAGCA 434
Qy 141 SerThrSerHisVal-----SerThrMetAlaGluProMetGluMet---TyrSerPro 157
Db 435 AGCACAAGTCAAGTGTCAATCTTAACACTGTGGACACAATGGATTTCCCACTCCGCA 494
Qy 158 ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAsp 177
Db 495 CCAACACACCAAGCAATATGATGCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 542
Qy 178 GlnSerSerCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSer 197
Db 543 -----GATCAATCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 581
Qy 198 TrpSerMetGlnLeuLeuAsnGlyAsp 206
Db 582 TGGTCTATGCAACTGCTAATGAGAG 608

```

RESULT 15

CV466307
 LOCUS CV466307 638 bp mRNA linear EST 30-SEP-2004
 DEFINITION Mdfrt3092e12.y2 Mdfrt Malus x domestica cDNA clone Mdfrt3092e12 5' similar to TR:Q39028 Q39028 ATMYB2.1, mRNA sequence.
 ACCESSION CV466307
 VERSION CV466307.1 GI:52860899

KEYWORDS
 SOURCE ORGANISM
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 .source

EST.
 Malus x domestica (cultivated apple)
 Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
 1 (bases 1 to 638)
 Korban,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A., Aldwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I., Tsagareishvili,R., Kennedy,S., Waterston,R. and Wilson,R.
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)
 Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library materials provided by: Schuyler S. Korban Library constructed by: K. Gasic Library sequenced by: Washington University Genome Sequencing Center
 WashU EST name: aal36c06.y2
 Seq primer: -40UP from Gibco.
 Location/Qualifiers
 1..638
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /db_xref="taxon:3750"
 /clone="Mdfrt3092e12"
 /lab_host="DH10B ampicillin resistant"
 /clone_lib="Mdfrt"
 /note="Vector: pBluescript II SK (+); Site_1: NotI; Site_2: EcoRI; Total RNA was extracted separately from each stage [young fruitlet(<1cm), young fruitlet (1 cm dia.), young fruitlet (12cm dia.), maturing fruit I, maturing fruit II, mature fruit], using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo18 (dT) primer with an identifying tag sequence (see table below). cDNA's from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (young fruitlet) insert 18(A)TGGTG; Stage 2 (young fruitlet 1cm dia) insert 18(A)TCGTG; Stage 3 (maturing fruit I) insert 18(A)TCGGA; Stage 5 (maturing fruit II) insert 18(A)TCGGA; Stage 6 (mature fruit) insert 18(A)TCGTG; Tag identification when sequencing from 3' end: Stage 1 (young fruitlet) CACGA18(T) insert; Stage 2 (young fruitlet 1cm dia) CACGA18(T) insert; Stage 3 (young fruitlet 12cm dia) ACCGA18(T) insert; Stage 4 (maturing fruit I) TCGCA18(T) insert; Stage 5 (maturing fruit II) TCGCA18(T) insert; Stage 6 (mature fruit) ACGCA18(T) insert. Double stranded cDNAs were size selected (more than 450 bp), adaptor with EcoRI adapters at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector(Stratagene). Identification of adaptors and tags in 5'-end sequenced clones:
 <Vector>...TAAGCTT-End Vector><Start
 EcoRI_adaptor>GATCGAATTCATGTTGTTGGG <End
 EcoRI_adaptor><Start Insert>...AAAAAAAAAAAAA-End
 Insert> <Start Tag>TCGCA<End Tag><Start
 NotI site/Vector>GCGGCCGCCACCGGG... The total number of white colony forming units (cfu) in the primary library before amplification was 2.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 4 kb, as determined by PCR. Purified plasmid DNA from the primary library was

converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 5.6×10^6 cfu. Background of empty clones was less than 1×10^6 .

ORIGIN

Alignment Scores:

Pred. No.:	9.23e-73	Length:	638
Score:	749.50	Matches:	149
Percent Similarity:	77.40%	Conservative:	12
Best Local Similarity:	71.63%	Mismatches:	30
Query Match:	65.86%	Indels:	17
DB:	7	Gaps:	5

US-10-659-869A-36 (1-206) x CV466307 (1-638)

```

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTirp 20
Db 57 ATGGATAAAATCCATTTCATAGTTCATCCAGGATGTTGAAGTCGAGAAAAGGCCCATGG 116

Qy 21 ThrMetGluGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTirp 40
Db 117 ACCATGGAAGAAGATTGATTCTCATCACTATATATGCAATCATGTTGAAGGTGTATGG 176

Qy 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTirp 60
Db 177 AACTCCCTAGCAAGACTGCTGGTCTCAACGTACTGGGAAGAGCTGCCGACTCCGGTGG 236

Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 237 CTGAATTATCTCGCGCTGATGTTTCGGAGAGGAAATATCACTCCTGAGGAACAACCTTTTG 296

Qy 81 IleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisLeuPro 100
Db 297 ATTATGGAACATGCAATGCAAGTGGGGTAACAGGTGGTCAAGAAATCGCAAAACATCTACCG 356

Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTirpArgThrArgIleGlnLysHisIleLys 120
Db 357 GGAAGGACTGATACGAATCAAGAACTATTGGAGGACTAGAAATCAAAGACACATTAAAG 416

Qy 121 GlnAlaGluAsnPheGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 417 CAAGCTGAGAAATATTACACCAGGACAAAGC-----TCTGAGGTCAATGAT---CAAGCA 467

Qy 141 SerThrSerHisVal-----SerThrMetAlaGluProMetGluMet---TyrSerPro 157
Db 468 AGCACAGTCAAGTGTCAATCTCTAACACTGTGGACACAATGGATATTTCCTCACTCCGCA 527

Qy 158 ProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAsp 177
Db 528 CCAACACACCAAGCAAGATATGATGCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 575

Qy 178 GlnSerSerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTirpSerMetGluAspSer 197
Db 576 -----GATCAATCTAATGATACTACTGAGCATGGAAGATCTC 614

Qy 198 TrpSerMetGlnLeuLeuAsnGly 205
Db 615 TGGTCTATGCAACTGCTAAATGGA 638

```

Search completed: September 28, 2005, 15:56:02
 Job time : 3169 secs

This Page Blank (uspto)